A Publication of the American Psychological Association

The Journal of Abnormal and Social Psychology

Founded by Morton Prince, 1906

CONTENTS

Articles	Page
An Experimental Comparison of Conclusion-Drawing by the Communicator and by the Audience *Carl I. Hovland and Wallace Mandell*	581
Time Orientation and Social Class *Lawrence LeShan*	589
Factors Related to the Institutional Adjustment of Prison Inmates *Patrick J. Driscoll*	598
Stability of the Self-Concept as a Dimension of Personality *John J. Brownfain*	597
The Rorschach Test and Suggestibility *Ira M. Steisel*	607
Conditioning and Extinction of the Galvanic Skin Response as a Function of Anxiety *M. E. Bitterman and Wayne H. Holtzman*	615
Acceptance by Others and Its Relation to Acceptance of Self and Others *Charles J. McIntyre*	624
Ratings on Traits Measured by a Factored Personality Inventory *John B. Carroll*	626
Stimulus Generalization After Extinction and Punishment: An Experimental Study of Displacement *F. R. Brush, R. R. Buth, W. O. Jenkins, W. F. Johns, and J. W. M. Whiting*	633
Rao's Multiple Discriminant Technique Applied to Three TAT Variables *Harold Webster*	643

Continued on inside front cover

Volume 47 JULY 1952 Number 3

CONTENTS—*Continued*

Verbal Attitudes and Overt Behavior Involving Racial Prejudice . . 649
Bernard Kutner, Carol Wilkins, and Penny Rechtman Yarrow

"Regression" in Schizophrenia as Determined by Performance on the Bender-Gestalt Test 653
Barbara J. Suttell and Gerald R. Pascal

The Discomfort-Relief Quotient as a Measure of Tension and Adjustment 658
Arnold Meadow, Milton Greenblatt, Julius Levine, and Harry C. Solomon

Intelligence from a Perceptual Point of View 662
Arthur W. Combs

A Comparison of Three Extinction Procedures Following Heart Rate Conditioning 674
J. M. Notterman, W. N. Schoenfeld, and P. J. Bersh

An Experimental Investigation of the Hypnotic Dream 678
Anders Sweetland and Herbert Quay

Aggression in Problem Children and Normals as Evaluated by the Rosenzweig P-F Study 683
Saul Rosenzweig and Louise Rosenzweig

Infiltration and the Attitudes of White and Negro Parents and Children 688
Charles Bird, Elio D. Monachesi, and Harvey Burdick

The Validity of Hypnotically Induced Color Hallucinations . . . 700
Bernard G. Rosenthal and Howard Mels

The Honi Phenomenon: A Case of Selective Perceptual Distortion . 705
Warren J. Wittreich

CASE REPORT

Working with a Psychopath 713
Louis C. Weber

CRITIQUES AND NOTES

On Keet's Study, "Two Verbal Techniques in a Miniature Counseling Situation" 722
Reed M. Merrill

The Effect of Anonymity on Attitude-Questionnaire Response . . 722
Philip Ash and Edward Abramson

Effects of Seating Arrangement on Leaderless Group Discussions . 724
Bernard M. Bass and Stanley Klubeck

The Role of Arbitrariness in the Frustration-Aggression Hypothesis . 728
Nicholas Pastore

Is the F Scale Indirect? 732
Thomas S. Cohn

Perceptual Defense or Response Suppression? 732
Edna M. Whittaker, J. C. Gilchrist, and Jean W. Fischer

REVIEWS 734

NEW BOOKS RECEIVED 745

Volume 47 JULY 1952 Number 3

The Journal of Abnormal and Social Psychology

Editor: J. McV. Hunt, University of Illinois

Book Review Editor: Stanley G. Estes, Harvard University

Managing Editor: Lorraine Bouthilet

ASSOCIATE EDITORS

Gordon W. Allport
Harvard University

Hadley Cantril
Princeton University

Lee J. Cronbach
University of Illinois

John Dollard
Yale University

Leonard W. Doob
Yale University

Charles M. Harsh
U. S. Navy Electronics Laboratory
San Diego

William A. Hunt
Northwestern University

Arthur Jenness
Williams College

W. Ernst Kris
Yale University School of Medicine

Donald W. MacKinnon
University of California

Quinn McNemar
Stanford University

O. Hobart Mowrer
University of Illinois

Theodore M. Newcomb
University of Michigan

Charles E. Osgood
University of Illinois

Saul Rosenzweig
Washington University

Harold Seashore
Psychological Corporation

David Shakow
University of Illinois School of Medicine

Ross Stagner
University of Illinois

Elizabeth Collins Wesman
White Plains, N.Y.

Robert W. White
Harvard University

PUBLISHED BY

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, Inc.

PUBLICATION OFFICE: 372 BROADWAY, ALBANY, N.Y., AND

1515 MASSACHUSETTS AVENUE, NW., WASHINGTON 5, D.C.

Entered as second-class matter, June 18, 1923, at the Post Office at Albany, New York, under the Act of February 28, 1925. Acceptance for mailing at special rate of postage provided for in Section 1103, Act of October 3, 1917, authorized August 26, 1947.

Copyright 1952 by the American Psychological Association, Inc.

The JOURNAL is issued quarterly in January, April, July, and October. The yearly volume averages 544 pages. The subscription price per volume is $6.00; foreign, $6.50. Single issues are $1.75.

Original manuscripts and editorial communications should be sent to:

DR. J. MCV. HUNT
Department of Psychology
University of Illinois
Urbana, Ill.

Books for review should be sent to:

DR. STANLEY G. ESTES
Psychological Clinic
64 Plympton Street
Cambridge 38, Mass.

Five copies of the JOURNAL are supplied gratis to contributors of articles and book reviews. Authors may order reprints when returning proof. No reprints are supplied gratis.

Printed and mailed by:

THE BOYD PRINTING COMPANY, INC.
372 Broadway, Albany 7, N.Y.

Subscriptions and business communications should be sent to:

THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
1515 Massachusetts Avenue, NW.
Washington 5, D.C.

AN EXPERIMENTAL COMPARISON OF CONCLUSION-DRAWING BY THE COMMUNICATOR AND BY THE AUDIENCE

CARL I. HOVLAND AND WALLACE MANDELL

Yale University [1]

IN preparing communications designed to influence opinion, the question can be raised as to whether it is more effective to present the factual information and arguments and let the audience draw its own conclusions or whether it is better for the communicator to draw the appropriate conclusion. Greater effectiveness of the former might be predicted from the assertion sometimes made that indirect suggestion is more effective than direct. A similar prediction might be made from the frequently cited tenet of the nondirective school of psychotherapy that decisions are more effective when reached independently by the client than when suggested by the therapist (4, p. 361). On the other hand, the opposite prediction can be made on the grounds that for many members of the audience it would be necessary for the conclusion to be explicitly stated to insure its being clearly perceived. It is also possible that the omission of an explicit statement of one's conclusion may serve as a cue that the communicator has something to conceal, and hence an explicit stating of the conclusion would be more effective. The present study was designed to study the effects upon beliefs of explicit stating of the conclusion under the conditions which typically obtain when information-education materials are presented through mass media. Experimental comparison was made of two types of communication which were identical in every way, except that in one the conclusion was drawn at the end by the communicator, while in the other the conclusion was not drawn by the communicator but left to the audience.

[1] This study was done as part of a program of research on factors affecting changes in opinion and attitude through communication, under a grant from the Rockefeller Foundation (see Hovland [1]). The writers wish to thank Prof. L. W. Crafts and the members of his teaching staff for their kindness in making their classes available for this study. They also wish to thank Drs. Irving L. Janis, Harold H. Kelley, and Fred D. Sheffield for their valuable suggestions given in the course of reading the manuscript.

One would expect that any generalization in this area would have to take into account the kind of communication, the nature of the communicator, and the type of audience to whom the communication is presented. It was thought that an important factor would be the degree of confidence which the audience has in the impartiality of the communicator. One hypothesis to be tested is that the drawing of the conclusion by the communicator will be less effective when his motives are suspected, and that here the individual's working through to the conclusion himself may make it more compelling. Accordingly, for each type of conclusion procedure two different communicators were used, one designed to arouse suspicion and the other confidence in his impartiality.

A second factor investigated was the intellectual ability of the members of the audience. It would be expected that the brighter members of the audience would have less need to have the implications of the premises spelled out, whereas the less intelligent individuals may be unable to arrive by themselves at the correct conclusion from the premises alone.

Finally, it was thought that the relative effectiveness of the two procedures might be dependent upon general personality traits of the audience members. It might be expected that some members of the audience would be resistant to suggestion and have to work through to their own conclusions, while others would be highly dependent upon the opinions of others and more inclined to follow the explicit conclusion of the communicator. To obtain information on this point, personality test items were included in the questionnaire administered before the communication.

Thus the basic problem investigated was the influence of explicit drawing of the conclusion by the communicator upon the audience's attitudes and beliefs in the area covered, but this relationship was studied as

a function of the interacting variables of (*a*) confidence in the communicator, (*b*) intellectual ability, and (*c*) personality traits possessed by members of the audience.

PROCEDURE

The over-all design of the study was to present communications which were identical except for the two experimental variations: the presence or absence of an explicit statement of the conclusion at the end, and an introduction which elicited either suspicion of the communicator's motives or belief in his impartiality. Thus four different versions of the communication were required. Opinions on the issue discussed in the communication were measured before the communication and again after the communication. In addition, a number of fact-quiz items were included in the questionnaire given afterwards, to determine the amount of information derived from the communication.

The Communication

The conclusion-drawing variable. Although numerous examples of this variable can be readily found, it is difficult to find the clear-cut extreme variation of the sort needed for a single-communication experiment. Particularly difficult to meet was the requirement that the communication present all of the basic material that would constitute the basis for drawing the conclusion but still not have the deduction obvious. A number of attempts were made to organize a communication along the lines of a syllogism in which the premises were presented and the conclusion was derived from the premises. Experience with this scheme was convincing evidence of the contention that syllogisms are essentially tautalogical. It proved extremely difficult to develop a communication where the basic premises were clearly stated that did not in the process "give away" the conclusion. It is possible that some more artificial communication would have succeeded along these lines but any one which seemed at all natural made the conclusion obvious. Some experimentation was also attempted using an inductive presentation in which the raw materials were presented on the basis of which the inductive generalization could constitute the conclusion, but it here also turned out to be difficult to present the material without organizing it in some way which revealed the objective behind the marshalling of the facts.

Finally, a method was devised which seemed to satisfy the major requirements of the design. The method employed involves a pattern similar to the "symptom and diagnosis" problem found in medicine. A communication was developed in which the general *principles* of the topic were presented together with a statement of the *conditions* existing, so that "conclusion-drawing" consisted essentially of fitting the particulars and the generalizations together to determine the appropriate implications.

The topic which was finally selected as lending itself well to this method was the topic "Devaluation of Currency." An introduction designed to arouse interest described briefly the history of our monetary system from the time of the early goldsmiths. The general principles underlying increasing or decreasing the metal equivalent of currency were then described, together with a statement of the conditions which make devaluation desirable or undesirable. Two historical examples were cited, one (Australia) illustrating the conditions in which devaluation is highly successful and another (Germany) in which the evils of inflation resulted because of the presence of conditions producing the opposite effect. Finally, the conditions currently obtaining in the United States were described. The drawing of the conclusion consisted of combining appropriately the general principles and the current conditions. A highly simplified economics was used in which the general effects of devaluation were given as increase in exports, decrease in imports, and increased employment. The conditions which would be favorable to devaluation were given as: need for foreign markets, a labor surplus, availability of goods for export, and a stable government. The corresponding factors suggesting avoidance of devaluation were given as: low productive capacity, a shortage of goods within the country, a labor shortage, and an unstable government. The conditions described as obtaining in the United States at the time of the experiment were slanted to indicate the following: the U. S. had a high productive capacity, a need for foreign markets, some excess in labor supply, and a very stable government. Thus the appropriate conclusion from the evidence presented would be the desirability of *devaluation* of American currency.

The "suspicion" variable. The communication was prepared on a transcription described as being taken off the air from a program called "Education for Americans." The speaker was accordingly introduced by the news commentator in charge of the program. This arrangement made it possible to have an identical communication but with different characterizations of the speaker. The recording was done on a tape so that one or the other introduction could be spliced to the same communication.

In one introduction an attempt was made to convey subtly that the speaker would have something to gain by having his conclusion followed, while in the other an impartial communicator was suggested. In the first version the announcer introduced the speaker as the head of a large importing firm. In the talk proper it was explicitly stated that importers would profit by devaluation of the dollar. In the alternative version the speaker was introduced as an economist from a specified leading American university. The intent was to arouse suspicion of the communicator's motives in the first version but to suggest impartiality in the second.

The Opinion Questionnaires

Two days before the communication a short questionnaire was administered ostensibly as a pretest of an information test which was being standardized for use by a nation-wide magazine. While the bulk of the questions dealt with general, political, and social issues, a number of key items were

inserted having to do with the subjects' opinions on the problem of devaluation. The questions were in four-alternative multiple choice patterns and even the opinion questions were in the context of information about current events.

The communications were presented as being a pretest by a commercial organization of a new type of current events program about which the opinion of the audience was sought. Following the playing of the tape recordings a questionnaire was administered dealing with audience reaction to the program. In this second questionnaire a number of opinion questions were asked which were identical to those used in the before questionnaire. A fact quiz was also included which was designed to determine the extent to which the audience had learned the basic content of the program, both the *principles* and the current *conditions* which together would determine the correct conclusion.

Personality Items

At the same session at which the first opinion questionnaire was given, a separate sheet was passed out containing 16 personality test items designed to distinguish between *Ss* who are dependent upon others for their opinions and those who like to form their own opinions. A sample question was: "Do you usually prefer to work things out for yourself than get someone else's advice?" The rationale given to *Ss* for this portion of the procedure was that data were needed on a national sample for purposes of standardizing a new personality test, and that to avoid any burden to the individual respondents the total questionnaire had been subdivided into a series of parallel forms.

Intelligence Test Scores

The majority of the *Ss* had already taken the ACE intelligence test in connection with the university's student personnel program. Accordingly, arrangements were made to enter each individual's score on the code sheet prepared to record responses to the questionnaires administered before and after the communication. The scores were subsequently subdivided into those above and those below the national norms on the ACE test.

Subjects

The *Ss* were 242 undergraduates at New York University in eight introductory sections of psychology. Each section contained between 25 and 40 students. The classes were assigned to the different treatments in such a way that there were two replications of each of the four major conditions.

RESULTS

The over-all results on opinion change for the four principal combinations of communicator and conclusion-drawing are given in Table 1. The data are derived from the responses to two key questions concerning the *S*'s evaluation of the desirability of America's devaluating the dollar. Answers to each question were dichotomized. On the first question (reproduced at the bottom of Table 1) the first two categories were scored as positive and the last three as negative. On the second question the first category was scored positive and the remaining two negative. Each *S* was then given a score which was −1 if responses to both questions were answered in the negative categories, 0 if one question was answered positively and the other negatively, or +1 if both questions were answered positively. The results are given in terms of the percentages of *Ss* who increased their scores (became more favorable) from before the communication to afterward, those who stayed the same, and those who decreased their scores. The "net change" score is computed by subtracting the unfavorable changes from the favorable.

Effects of Conclusion-Drawing on Opinion

Comparison of the two conditions in which the speaker drew the conclusion with the other two conditions in which he did not draw the conclusion reveals a very large difference, in favor of conclusion-drawing by the communicator. Over twice as many *Ss* changed their opinions in the direction advocated by the communicator when the conclusion was explicitly drawn than when left to the audience. The difference is significant at the .001 level when computed by the formula for net change discussed in Hovland, Lumsdaine, and Sheffield (2).

Effects of Suspicion-Arousing Introduction

The other primary variable used in the study was the arousal of suspicion of the communicator. As described under *Procedure,* suggestion of possible bias on the part of the communicator was introduced by giving one of the communicators a role in which he might have something to gain by the audience's accepting his position. The effects of this variable on opinion can be seen from Table 1. The difference in effect in favor of the nonsuspicious communicator is 7.2 per cent with the conclusion drawn by the communicator and 4.9 per cent with the conclusion not drawn by the commentator. The p value for the first is only .23 and for the second only .33. So, although both differences are in the predicted direction and the

difference obtains for both questions and for both of the replication subgroups, the probability value is not at an adequate level of significance.

The small difference in the effect of the suspicion variable raises the question as to whether the procedure for arousing suspicion was too subtle and whether the two communicators were not perceived as differing in trustworthiness. Information on this point can be derived from the answers to the questions asking each group to evaluate the competence, fairness, and honesty of each presentation. These data are presented in Table 2.

The results indicate that the suspicion-arousing introduction did cause the audience to view the "motivated" communicator as having done a less good job and as being less fair and honest in his presentation than the "impartial" communicator *despite the fact that the conclusions were identical.*

TABLE 1

PER CENT OF SUBJECTS CHANGING OPINION ON "DEVALUATION" FROM BEFORE TO AFTER COMMUNICATION *

	DIRECTION OF CHANGE **	CONCLUSION NOT DRAWN BY COMMUNICATOR %		CONCLUSION DRAWN BY COMMUNICATOR %	
"Nonsuspicious" Communicator	Positive	30.9		53.4	
	No Change	60.0		44.8	
	Negative	9.1		1.8	
	Net % Changing		21.8 (N=55)		51.6 (N=58)
"Suspicious" Communicator	Positive	30.5		49.2	
	No Change	55.9		46.0	
	Negative	13.6		4.8	
	Net % Changing		16.9 (N=59)		44.4 (N=63)
Combined ("Suspicious" and "Nonsuspicious" Communicators)	Positive	30.7		51.2	
	No Change	57.9		45.5	
	Negative	11.4		3.3	
	Net % Changing		19.3 (N=114)		47.9 (N=121)
Mean Diff.=			28.6		
p=			.001		

* Data based on changes in responses to the following two questions administered before and after communication:

Question 1. Devaluation of the dollar as a method for manipulating the American economy should be used: (Check one)

\+ a. ——— In a situation such as exists at present
\+ b. ——— If the situation starts to become worse
— c. ——— Only if things become very much worse
— d. ——— Only in an extreme emergency
— e. ——— Never

Question 2. In view of the current economic situation do *you* think the U.S. should: (Check one)

\+ a. ——— Devaluate its currency
— b. ——— Not change its currency value
— c. ——— Raise the value of its currency

** Positive is here defined as changing in direction of position advocated by communicator, negative as changing in direction opposite that advocated by communicator. Net % changing=per cent changing in positive direction minus per cent changing in negative direction.

Included in the second questionnaire were a number of fact-quiz items concerning points covered in the communication. It is of incidental interest to determine whether the suspicion of the communicator affected the amount of factual information learned from the communication. In Table 3 data are given for the mean number of items correct on the 12-item fact quiz concerning the eco-

TABLE 2

EVALUATIONS OF IDENTICAL COMMUNICATIONS WHEN PRESENTED BY "SUSPICIOUS" AND "NONSUSPICIOUS" COMMUNICATORS

	"NONSUSPICIOUS" COMMUNICATOR (N=113)		"SUSPICIOUS" COMMUNICATOR (N=122)
A. Question: "Do you think this radio program did a good or a poor job of giving the facts on devaluation of currency?"			
Per cent of subjects answering "A very good job"	41.1		21.1
Difference		20.0	
p		<.001	
B. Question: "Do you feel that the speaker was fair and honest about America's devaluating its currency or did the facts given seem too one-sided?"			
Per cent of subjects saying communicator gave a "fair and honest" picture	52.7		36.7
Difference		16.0	
p		<.01	

nomics of devaluation. There is no evidence that the suspicion-arousing introduction affected adversely the amount of information learned from the talk ($p = .34$).

Interaction of Suspicion-Arousal and Conclusion-Drawing

The hypothesis that conclusion-drawing by the audience may be more effective when the communicator arouses suspicion was tested by comparing the effect on opinion of conclusion-drawing by the communicator and by the audience with the "suspicion" and "no-suspicion" introductions. From the data presented in Table 1 it is found that there is a differential change of 27.5% for the suspicious communicator and 29.8% for the nonsuspicious so there is no evidence of any differential effect of conclusion-drawing under the two conditions.

From Table 1 it will be clearly seen that the conclusion-drawing and suspicion factors operate cumulatively and that the "no-suspicion" "conclusion-drawing" condition produces a difference of 51.6% while the "suspicion" "no conclusion-drawing" condition produces a difference of only 16.9%. The difference of 34.7 is significant at less than the .001 level.

Effects for Groups of Different Intellectual Ability

It was initially thought that individuals of high intellectual ability would be little affected by the conclusion-drawing variation, but that in the case of individuals of lower ability the drawing of the conclusion by the communicator would be effective because it would guarantee the understanding of the implication of the premises presented by the communicator.

Accordingly, *Ss* were divided into two subgroups in terms of intelligence test scores.

TABLE 3

MEAN NUMBER OF FACTUAL ITEMS CORRECT FOLLOWING COMMUNICATION WITH "SUSPICIOUS" AND "NONSUSPICIOUS" COMMUNICATORS PRESENTING IDENTICAL MATERIAL

	"NONSUSPICIOUS" COMMUNICATOR (N=112)		"SUSPICIOUS" COMMUNICATOR (N=122)
Mean number of items correct	9.61		9.70
Difference		0.09	
p		<.34	

TABLE 4

COMPARISON OF NET PER CENT OF SUBJECTS CHANGING OPINION IN DIRECTION OF COMMUNICATION AMONG HIGH AND LOW INTELLIGENCE SUBGROUPS WHEN CONCLUSION IS DRAWN OR IS NOT DRAWN BY SPEAKER *

	DIRECTION OF CHANGE	PER CENT OF INDIVIDUALS CHANGING OPINION					
		HIGHEST INTELLIGENCE			LOWEST INTELLIGENCE		
Conclusion *Drawn* by Communicator	1. Positive	45.3		1. Positive	52.0		
	2. No Change	54.7		2. No Change	40.0		
	3. Negative	0.0		3. Negative	8.0		
	Net Per Cent Changing		45.3 (N=53)			44.0 (N=25)	
Conclusion *Not Drawn* by Communicator	1. Positive	23.9		1. Positive	32.1		
	2. No Change	63.1		2. No Change	53.6		
	3. Negative	13.0		3. Negative	14.3		
	Net Per Cent Changing		10.9 (N=46)			17.8 (N=28)	

* Questions same as those in Table 1.

The cutting point was the median score on the national norms for college students provided for the ACE examination. Approximately two-thirds of the *S*s upon whom we had test information scored above the national average. The amount of opinion shift in the direction of the communication with the "conclusion-drawing" and "no-conclusion" versions for the "high" and "low" intelligence group is presented in Table 4.

The results indicate that within the narrow range of intelligence scores represented at this university there is no relationship between intelligence and the influence of conclusion drawing.

Personality Factors

It was initially thought possible that individuals who indicated on personality questionnaires that they liked to make up their minds for themselves and were resistant to suggestions by others might be less influenced by the communication in which the conclusion was completely spelled out by the communicator than the one in which the conclusion-drawing was left up to them. Results for the third of the audience giving most "resistant" answers in the personality questionnaire compared with the third giving most "suggestible" answers are presented in Table 5.

TABLE 5

COMPARISON OF NET PER CENT OF INDIVIDUALS CHANGING OPINION IN DIRECTION OF COMMUNICATION WITHIN SUBGROUPS OF "RESISTANT" AND "SUGGESTIBLE" INDIVIDUALS WHEN CONCLUSION IS DRAWN OR IS NOT DRAWN BY COMMUNICATOR *

	DIRECTION OF CHANGE	PER CENT OF INDIVIDUALS CHANGING OPINION					
		"RESISTANT" INDIVIDUALS			"SUGGESTIBLE" INDIVIDUALS		
Conclusion *Drawn* by Communicator	1. Positive	62.9		1. Positive	51.4		
	2. No Change	34.2		2. No Change	42.9		
	3. Negative	2.9		3. Negative	5.7		
	Net Per Cent Changing		60.0 (N=35)			45.7 (N=35)	
Conclusion *Not Drawn* by Communicator	1. Positive	27.3		1. Positive	29.4		
	2. No Change	66.6		2. No Change	58.8		
	3. Negative	6.1		3. Negative	11.8		
	Net Per Cent Changing		21.2 (N=33)			17.6 (N=34)	

* Questions same as those in Table 1.

No difference is obtained in the effects of conclusion drawing for the two groups. But whether the personality questionnaire is itself a valid measure is brought into question by the absence of the expected positive relationship between "suggestibility" scores and opinion change. The nonsignificant relationship which is obtained is in the negative direction, i.e., the more suggestible on the personality test being less influenced by all communications.

DISCUSSION

The operation of the "suspicion of communicator" variable in the present experiment appears to parallel closely the effects of the source-credibility variation in the experiment of Hovland and Weiss (3). They found that information contained in a communication was equally well learned whether the source was one considered to be "trustworthy" or one considered "untrustworthy," but that the effectiveness of the trustworthy source in affecting opinion was greater than that of the untrustworthy source. It is likely that in their experiment more extreme variation in the credibility of source was obtained and this may account for the smaller effects of the "suspicion" variable on opinion change in the present study.

The present results seem to be fairly clear-cut in showing a greater effect on opinion change when the conclusion is explicitly drawn by the communicator as compared with the change when the conclusion has to be drawn by the audience. This result may have considerable significance for communications similar to the one here employed. But a great deal of further work is needed to determine the generality of the conditions for this outcome. Only a single population was employed—a group with more than average intelligence and education. It seems likely that the difference would be even more pronounced with a wider range of education and intelligence. More serious is the fact that only a single issue was employed. The present issue was moderately complex and the role of conclusion-drawing may be more significant here than on issues where the conclusion is more obvious. Interesting further research is suggested with other kinds of issues, particularly ones with a higher degree of ego-involvements. Such issues may permit the operation to a greater extent of some of the factors which are stressed by psychotherapists in suggesting nondirective procedures.

Much further information is required on the factors responsible for the superiority of conclusion-drawing in situations like the one employed in the present study. It appears likely that the result is due in part to the degree of clarity of the conclusion in the minds of the hearers when it is explicitly drawn. This factor should operate particularly among individuals of limited intellectual ability, and among those who do not pay close enough attention to the material to see the interrelationship between the data presented.

One hypothesis mentioned in the introduction for the greater effectiveness of conclusion-drawing was that failure to explicitly state one's conclusion may constitute a cue for the audience that the communicator has something to conceal. No support for this hypothesis was furnished by analysis of two audience-evaluation questions included in the questionnaire given after the communication. One question concerned the number of individuals who thought the speaker one-sided while the second was concerned with the number of individuals who thought the speaker was trying to propagandize the audience. In the case of the former question there were actually more individuals who thought the speaker one-sided when the conclusion was explicitly drawn. For the latter question there was no difference in attribution of propagandistic attempt to the speaker under the conclusion-drawing and no-conclusion conditions. Thus, even though more people thought the speaker one-sided when he stated his position, more individuals changed in the direction advocated by the speaker when he did make his position clear. It is possible that the speaker's having a position to defend may have had less effect on the opinions of the audience in the present situation than in others that can be conceived where the impartiality of the communicator is extremely critical.

In the present experiment the analysis of effects of conclusion-drawing were limited to short-term changes in opinion. It would be of considerable interest to investigate other

possible measures of effect. Would there be differences in retention? It is possible that when one arrives at his decision independently he holds it more persistently because of a higher degree of ego-involvement in the process. Would there be a difference in the likelihood that the conclusion reached independently would be more likely to be applied in other relevant contexts? Here tests of the effects of conclusion drawing on other types of behavior than opinion change would be required. Finally, there is an interesting possibility that when the conclusion is drawn for the audience by the communicator the effects on the specific issue may be greater, but that when one reaches the conclusion by himself he is more "sophisticated" about implications not covered by the communicator and that this difference would affect his deriving new implications.

SUMMARY

1. The effects upon opinion change of having the communicator draw the appropriate conclusion from material he had presented were compared with those produced when the drawing of the conclusion was left to the audience. For each type of presentation two communicators were used: one who would appear to have something to gain by having the conclusion accepted, the other an impartial individual. The topic presented was "Devaluation of Currency." The communications were transcribed on tape and were presented as radio programs with respect to which the experimenters wished the audience's comments and evaluations. Opinion questionnaires were administered before and after the communication. Personality questionnaire and intelligence test data were also secured. Eight classes of college students provided two replications of each treatment and gave a total sample of 242 students of both sexes.

2. Over twice as many *Ss* changed their opinions in the direction of the position advocated by the communication when the speaker drew the appropriate conclusion than when the drawing of the conclusion was left to the audience ($p = .001$).

3. The communicator introduced as belonging to a group which would have a motive to convince the audience of his position was judged to have given a more one-sided presentation of the facts than the "impartial" communicator, even though the talk itself was identical.

4. The amount of factual information acquired was not significantly different for the "suspicion-arousing" and "impartial" communicator versions. This supports the results of Hovland and Weiss (3).

5. The number of individuals who changed their opinions in the direction of the communicator's position was in the predicted direction of being greater for the "impartial" than for the "suspicion-arousing" communicator version, but not at an adequate level of significance.

6. The differential effects of conclusion-drawing by the communicator versus by the audience were about the same whether a "suspicion-arousing" or an "impartial" communicator was employed. Thus there was no support for the hypothesis that drawing one's own conclusion is particularly effective when one suspects the motives of the communicator.

7. The combined operation of the variables of suspicion and conclusion-drawing produced a highly significant effect, the unbiased communicator who drew the conclusion being the most effective and the suspicion-arousing communicator who did not draw the conclusion being least effective.

8. No significant relationship was obtained between responses to the personality test items designed to measure "resistance to suggestion" and "independence" and the influence of the conclusion-drawing variable.

9. No relationship was obtained between intelligence test score and the efficacy of conclusion-drawing. This may have been due to the narrow range of intelligence represented in a selected university population.

REFERENCES

1. HOVLAND, C. I. Changes in attitude through communication. *J. abnorm. soc. Psychol.*, 1951, **46**, 424–437.
2. HOVLAND, C. I., LUMSDAINE, A. A., & SHEFFIELD, F. D. *Experiments on mass communication.* Princeton: Princeton Univer. Press, 1949.
3. HOVLAND, C. I., & WEISS, W. The influence of source credibility on communication effectiveness. *Publ. Opin. Quart.*, 1951, **15**, 635–650.
4. ROGERS, C. R. Some observations on the organization of personality. *Amer. Psychologist*, 1947, **2**, 358–368.

Received October 22, 1951.

TIME ORIENTATION AND SOCIAL CLASS

LAWRENCE L. LESHAN
Roosevelt College

One of the most fruitful areas of research in modern sociology has been the investigation of the various levels of social class in America. An extensive literature on this subject has arisen in the past ten years. Volumes such as Havighurst and Taba's *Adolescent Character and Personality* (7), Gardner and Davis' *Deep South* (4), and Warner, Meeker, and Eell's *Social Class in America* (10) have summed up some of the major studies.[1] Out of the vast amount of data collected, many of the important implications have been extracted and reported. However, in as rich and new a field as this, there are always new implications to be explored and tested.

This paper is an attempt to examine one of the possible differences between members of social classes in America. This variable might be termed "collective-ego-space-time" after Erickson (5). It concerns the perceived relationship of the individual and his goals in time. How far ahead is the time span with which the individual is concerned? What is the crucial time limit during which he will frustrate himself in order to attain a goal? Does he relate his behavior primarily to the far future, the immediate future, the present, or the past? Bateson has pointed out that this orientation seems to differ widely between Balinese culture and the culture of the (middle class) United States. There would be important implications in knowing if it differed in various classes within the United States itself.

The Hypothesis

The general hypothesis of this paper is that there are various temporal goal orientations in the various levels of social class. Very briefly, these temporal orientations may be described as follows:

1. In the lower-lower class, the orientation is one of quick sequences of tension and relief. One does not frustrate oneself for long periods or plan action with goals far in the future. The future generally is an indefinite, vague, diffuse region and its rewards and punishments are too uncertain to have much motivating value. In this social class, one eats when he is hungry; there are no regular meal hours and each member of the family takes food when he feels like it if food is available.

2. In the upper-lower, middle, and lower-upper classes, the orientation is one of much longer tension-relief sequences. As the individual grows older, he plans further and further into the future and acts on these plans. As an adult, he may start planning for retirement when he is in his twenties. In these classes, one eats at regular "clock" hours. One quickly learns to inhibit activity leading to the relief of a basic tension (food-getting behavior) until a watch shows that it is time to eat.

3. In the upper-upper class, the individual sees himself as part of a sequence of several or more generations, and the orientation is backward to the past. One eats at traditional hours and lives out the traditions set up in the past.

This hypothesis has been implied in many of the books and articles on social class. A few quotations may serve to demonstrate this.

> . . . in a lower class family . . . the disciplinarian . . . is certain to believe that the way to make a child learn is to beat him. . . . It seems clear . . . that a child cannot be trained in this fashion to undergo the long periods of renunciation which the middle class ideal of socialization demands of him (2, p. 267).

> Every individual in the upper class has a series of duties and privileges associated with . . . his direct lineal ancestors, symbols of the past (4, p. 84).

> The greatest possible insult to an upper-upper class person is the defamation of his "original ancestor," the founder of his local line (4, p. 85).

> Both upper-middle-class and lower-middle-class husbands and wives recognize the raising of a family as the primary function of their relationship, more important than the sexual enjoyment of their partners, their economic security or their general physical comfort (4, p. 100).

[1] Clear definitions of the various social classes may be found in any of the above-mentioned volumes. In the interest of space saving, these definitions are not included here.

Evidences from Child-Rearing Practices

We would expect these various temporal orientations to be demonstrated in the techniques with which parents from these social classes train children. We would expect this for the double reason (*a*) that the parents themselves possess this orientation and (*b*) that we believe that children will, by and large, also have it when they mature. In studying training techniques as revealed in the various books and articles on social class, we find this expectation confirmed. In the third group (upper-upper) as postulated in this paper, we find training techniques such as: "What would your grandmother say?" or "Your grandfather would rather see you dead" (4, p. 98). In the second group (upper-middle, middle, and upper-lower), stress is on the future. Children are exhorted to do well in school, for example, by threats that they will not be able to get a good job, that they will not get a good spouse, or that they will not get into college (8). In the first group (lower-lower), training techniques are more in terms of immediate punishment and reward. Children are made to do (or stop doing) things by the threat that something will happen to them immediately if they do not obey (3, 4, 8). In this class, thumb-sucking and masturbation are stopped by threats of an immediate beating, physical punishment, or by physical restraint (3, p. 18). A reference to moral or developmental reasons for doing or not doing something is rarely if ever made.

In one lower-lower case studied (3, p. 44) a ten-year-old son has "found" some ice skates and brought them home. The father looks at them and says: "Now, listen, Johnny, I don't want you keepin' them if there's going to be any trouble with them."

This same case history of a lower-lower family was analyzed for the methods that the parents used to control the children. The interviewer became a close friend and confidant of the family and spent a great deal of time at their home. She observed very frequent attempts of the parents directly to make the child do, or not do, something. These can be broken down into categories as can be seen in Table 1. This table seems to indicate an orientation in the parents such as we have postulated, which one would suspect would be likely to produce a similar orientation in the children. Also of interest is the fact that 28 of the parental responses were unpredictable from the point of view of the child—as well as the point of view of the interviewer! As an example, we mention the following incident: The father had brought a three-year-old daughter some candy cigarettes. The child pretended to smoke them, flicking off the ashes, etc. When she did this, the father took them away from her and spanked her for "putting on the dog."

Children of the lower-lower class soon learn that major changes in their lives often occur suddenly and unpredictably. The mother in the case history mentioned above describes the weaning of one child: "Yeah,

TABLE 1

Parental Control Methods in a Lower-Lower Class Family, an Analysis of the Items in an Unpublished Case Study (3)

Type of Control	Example	Number of Attempts Observed
1. Flat declarative statements	"Put down that bottle."	35
2. Immediate punishment or reward threatened	"You go back to bed or I'll beat you with this leather strap."	57
3. Reward or punishment to follow within 24 hours	(Taking a dirty toy from a child) "Wait until Tootsie comes home and we'll wash it and you can play with it."	7
4. Reward or punishment to follow within the week	(2 days before Xmas) "Santa Claus won't come if you're bad."	1
5. Reward or punishment after 1 week or at an indefinite time	"I'll never take you any place again."	1

one day she threw the bottle out of her crib and I gave it back to her and she didn't want it so I never gave it to her again" (3, p. 13). This is a far cry indeed from the careful, long-term weaning of the middle classes. With the lower-lower type of experience one might speculate that children would grow up not at all certain of the basic stability of the universe.[2]

Food, shelter, heat in the winter (and even personal safety) are also unpredictable. Food is present and eaten when father gets a pay check *if* he brings it home. There is never enough over a long period of time and one can never tell whether or not there will be food when he is next hungry.

In the lower-lower group there seems to be a circular phenomenon at work. The parents' training is inconsistent because of their inability to work for long-range goals, and this prevents them from breaking out of the economic trap they are in. Economic pressures further decrease stability. Children who go through this training will also emerge unable to work for long-range goals and so on.

Other factors add to the inconsistency of training of children in the lower-lower class in relation to that of children in the middle and upper-upper classes. The lower-lower class child plays on the streets away from adult supervision. Here he is to a large degree at the mercy of his own impulses with reward or punishment following immediately on his actions. Rewards here may be in terms of motor activity, physical gratification, etc. The parents are at work or are usually unaware of what he is doing.

In the other social classes, the play situation tends to be quite different. Children play near the house or near the house of a neighbor and are, a large percentage of the time, under some adult supervision. Less often do both parents work and there is much more watching and controlling the children's play. Figuratively, they tend to play "in front of the house." The child has more training in controlling his own impulses because of the help in control he receives from older figures.

According to psychoanalytic theory, the individual must have a strong superego if he is to frustrate himself and renounce present pleasure for future gains. This is built up out of the child's image of the parents in his early years. It rewards him for self-control and punishes him for transgressions. In the second and third groups set up in this paper, the parents and families tend to remain constant for years at a time. In the first group, families are much more flexible: adults move in with other adults; children may be switched from one family to another for a while if this arrangement seems to be more convenient. Under this situation, it would be improbable that a strong and stable ego-ideal could be built up. There would be many shifting adults, all too inconsistent in their behavior for the child to build a model.

[2] With such a frame of reference, planning for a distant future would not seem to be an intelligent procedure.

PROCEDURE

To test partially this time orientation hypothesis, the stories of 117 children of 8 to 10 years old were examined. These stories were told in response to the stimulus "Tell me a story." Seventy-four of the children were lower class and 43 were middle class. The groups were equated for age. The stories were examined in terms of the period of time covered by the action of the story. In line with the hypothesis, one might expect that the middle-class group would produce stories covering a longer time-period from beginning (the starting of action) to the end (final action) than would the lower-class group. This expectation was confirmed statistically (Table 2). A chi-square test indicated that the two groups were not from the same population insofar as this variable was concerned. There is a statistically significant difference between the length of time covered in stories told by children of these two social classes. If one accepts the hypothesis that in unstructured situations of this sort, individuals tend to project the world as they see it onto the stimulus, then our major hypothesis of different time orientations in different social classes is strengthened.

DISCUSSION

If our hypothesis of different time orientations for different social classes is valid, there are various implications which may be of importance, such as the need to think through again the problem of goals and methods in the public schools, and re-education in old age when a time span which includes plan-

ning for the future may lose its meaningfulness. It appears worth while to consider two such implications in a little more detail.

One of the trickiest problems in modern psychopathology has been the concept of the psychopath. Whether or not psychopaths exist, and if so what the symptoms and etiology are, has been discussed in a great many articles with no real agreement or even pragmatically useful results. The essence of the various definitions advanced seems to rest on two general symptoms: low frustration tolerance and marked hostility turned inward and/or outward. It may be of value to examine the psychopath category in terms of personal time orientation as discussed above. One might speculate that an individual raised in an environment where (*a*) reward and punishment generally follow immediately on action and (*b*) where these results are unpredictable a large part of the time, would have a low frustration tolerance. He would not have learned to act in terms of future reward and would, indeed, have learned the opposite since the future would be an unpredictable region and to work in terms of it would be nonsensical for him. Further, this orientation in a world that is primarily run on longer sequences might well produce conflicts, failures, and resulting hostility. This training factor has not, to the author's knowledge, been investigated and it may be that here lies one of the roots of psychopathy.

A further implication of the concept of personal time orientation is in the general area of the prevention of delinquency. A reform school will be neither a deterrent in the future nor a lesson from the past in an individual who has learned to respond only in terms of what is immediately present. It may well be that in order to control delinquency it will be necessary first to change the time orientations of the delinquents. At present, very little is known concerning the management of such learning and it would appear to be a fruitful subject for research.

Conclusion

The hypothesis that there are different personal time orientations in different social classes is advanced. Data are presented (*a*) from child-rearing practices in the various classes of American society and (*b*) from the greater time span of the action in stories told by middle-class children as compared with those told by lower-class children which tend to confirm this hypothesis. Some further implications of the hypothesis are discussed briefly.

TABLE 2

Length of Time Covered by Action of Stories Told by 117 Children In Response to Stimulus "Tell Me A Story"

Class	Under 1 Hr.	1–12 Hrs.	12–24 Hrs.	1 Day–7 Days	1–2 Wks.	2 Wks.–1 Yr.	1 Yr. Plus
74 Lower-Class Children	44% (32)	32% (24)	5% (4)	8% (6)	5% (4)	4% (3)	1% (1)
43 Middle-Class Children	28% (12)	23% (10)	9% (4)	23% (10)	9% (4)	2% (1)	5% (2)
	(df=6)				(p=.001)		

References

1. Bateson, G. Cultural determinants of personality. In J. McV. Hunt (Ed.), *Personality and the behavior disorders*, Vol. 2. New York: Ronald, 1944. Pp. 714–735.
2. Davis, A., & Dollard, J. *Children of bondage*. Washington: American Council of Education, 1940.
3. Davis, A., & Erickson, M. An unpublished case study. The Committee on Human Development, Univer. of Chicago.
4. Davis, A., Gardner, B. B., & Gardner, M. R. *Deep South*. Chicago: Univer. of Chicago Press, 1941.
5. Erickson, H. Childhood and tradition in two American Indian tribes. In C. Kluckhohn & H. A. Murray (Eds.), *Personality*. New York: Knopf, 1949.
6. Erickson, Martha. Child rearing and social status. *Amer. J. Sociol.*, 1946, 52, 3.
7. Havighurst, R., & Taba, Hilda. *Adolescent character and personality*. New York: Wiley, 1949.
8. Holingshead, A. B. *Elmtown's youth, the impact of social classes on adolescents*. New York: Wiley, 1949.
9. Mowrer, O. H., & Ullman, A. D. Time as a determinant in integrative learning. *Psychol. Rev.*, 1945, 52, 61–90.
10. Warner, L., Meeker, Marchia, & Eells, K. *Social class in America*. Chicago: Science Research Associates, 1949.

Received July 30, 1951.

FACTORS RELATED TO THE INSTITUTIONAL ADJUSTMENT OF PRISON INMATES [1]

PATRICK J. DRISCOLL
University of Wisconsin

Because of the nature of the institutional environment, special problems arise in predicting and understanding adjustment of the convicted inmate. This study, which attempts to disclose the nature of some of these special problems of inmate appraisal, describes a technique designed to obtain an over-all estimation of the adjustment of prison inmates. Research with this technique in a large state prison appears to clarify in certain ways understanding of inmate characteristics and to define several problems of prognosis of adjustment of convicts to prison life.

In the present study, an assay of personality adjustment in prison inmates, made by means of a special rating system, has been compared with appraisals made by means of clinical-type personality tests. The effort has been made to investigate the possibility that the special environment of the prison may produce demands on the individual which favor modes of conduct objectively appraised, in terms of standards of the general social environment, as essentially maladaptive in character. The investigation of such institutional effects on the personality of the individual represents an initial step necessary for the development of adequate procedures in prediction of prison adjustment.

Method

There have been few attempts made to develop devices suitable for measuring the general behavior of the prison inmate. In order to obtain such an appraisal, a rating method has been developed in this study which embodies the principles of analysis of concrete behavior rather than general trait characteristics.

The rating method developed is based on the assumption that adjustment will consist of the particular properties of conduct associated with four main areas of response: social behavior, vocational behavior, attitudes and activity oriented toward self, and general behavioral organization and response. In order to sample these four modes of conduct, a rating scale has been designed consisting of 16 behavioral factors, divided into four subgroups designated in terms of the four areas of behavior mentioned.

For practical purposes the scale developed has been called "Adjustment Rating Scale." It is headed with identification data and instructions for use. At the left margin of the scale are listed the 16 factors divided into the four subgroups as follows:

Social Factors

- How has he acted toward others?
- What is the nature of his social behavior?
- How mature is the man in his social relations?
- What are this man's social attitudes?

Vocational Factors

- Has he been interested in his work?
- Has he produced on the job?
- Has he sought to improve himself?
- How has he reacted to responsibilities?

Personal Factors

- Has he been stable emotionally?
- Has he confidence in himself?
- What has been his outlook on life?
- Has he understood himself?

Behavioral Factors

- Has he shown evidence of neurotic disorders?
- Has he been in good physical condition?
- Has he shown himself to be a complainer?
- Has he been energetic in speech and action?

To the right of each of the individual factors there are listed a negative and positive statement which in effect dichotomized the rating made on this factor. A number of identifying words and phrases are printed under each of these statements to clarify their meaning. Ratings are made on the basis of certainty of the statement's applicability to the inmate, with a range on an eight-point scale from complete certainty of the positive statement's applicability to complete certainty of the negative statement's applicability to the inmate. Perforation of printed dots provides a means of recording the rating and a numerical index of the rating can be secured from the back of the rating sheet.[2]

After the preliminary form of this rating system

[1] This research was conducted as a part of a general vocational research program conducted in Waupun State Prison by the Bureau of Industrial Psychology, University of Wisconsin.

[2] For copies of the Adjustment Analysis Rating Scale, order Document 3613 from American Documentation Institute, 1719 N St., N.W., Washington 6, D. C., remitting $1.00 for 35 mm. microfilm or $1.00 for photocopies (6 x 8 inches).

was devised, a training program in the use of the rating scale was initiated for prison work supervisors. When this period of training was completed, a preliminary study of the rating procedure was conducted on 138 inmates. Thereafter, the scale was refined and improvements made in the procedure.

The methods of the main study involved the following steps: (*a*) investigation of the reliability of the Adjustment Rating Analysis on the basis of supervisory ratings of 60 inmates; (*b*) comparison of adjustment ratings with number of bad conduct reports recorded for each of the 60 prisoners; (*c*) evaluation of the validity of two personality tests, administered to each of the prisoners, in predicting the adjustment ratings of the prisoners.

The determination of the reliability of the adjustment rating was based on a rating-rerating procedure. The 60 inmates used in this main part of the study were a different subject group from the 138 subjects used in the initial preliminary investigation of the rating system. Work supervisors in different industrial departments of the prison, who had been trained in the initial study, rated the inmate subjects who were assigned to their departments.

Bad conduct reports, concerning infractions of prison discipline, were recorded for each prisoner as a part of the regular procedure of prison administration. Most of the prisoners used in the study had been incarcerated for a period of about one year when this study was carried out, so a significant number of bad conduct reports were obtainable on many prisoners.

In order to study the problem of prediction of prison adjustment, as determined by the ratings, two personality tests were administered to the subject inmates. One of these tests was the Minnesota Multiphasic Personality Inventory (MMPI) (3), which was given when the inmates first entered prison. All inmates used in the study had validating scores (?, L, F) on this test not exceeding 55, which is approximately at the center of the normal range. The results on the MMPI, which was administered as a part of a regular vocational appraisal of incoming inmates, were compared with the over-all adjustment ratings on the inmates.

A second personality test was also used in the study. This test is a "Prison Projective Test" developed especially for this study.[8] The test consists of 10 pictures, modelled in general after the Thematic Apperception Test, which depict unstructured prison situations. Inasmuch as the results of this test showed no significant relations with the data obtained from the adjustment ratings, the findings obtained with the test will not be discussed further in this report.

RESULTS

The results of the study reported here deal with three main points: (*a*) reliability of the rating scale developed, (*b*) the relation between the rated adjustment of prison inmates and other measurable general characteristics of the convicts, and (*c*) the correlation between personality test scores obtained at the time of initial incarceration and the later rated adjustment of the inmate.

Reliability of the Adjustment Analysis Rating Scale was determined by the rating-rerating method, with approximately one and one-half months intervening between ratings. Reliability coefficients for the various categories are +0.96 for the social category, +0.92 for personal category, +0.80 for vocational category, and +0.74 for the behavioral category of ratings. The rating-rerating reliability of the total scale is +0.85.

A first question of significance for this study was whether rated adjustment of the inmate corresponds with other observable aspects of adjustment in prison. The 60 inmates were divided into four groups of 15 each on the basis of the over-all ratings. The lowest group, Group I, received the poorest ratings, and the high group, Group IV, received ratings indicating the most successful adjustment. The number of bad conduct reports received by the various groups during a period of about one year in prison was then determined. Figure 1 illustrates the difference in the number of bad conduct reports in the various groups. The differences between

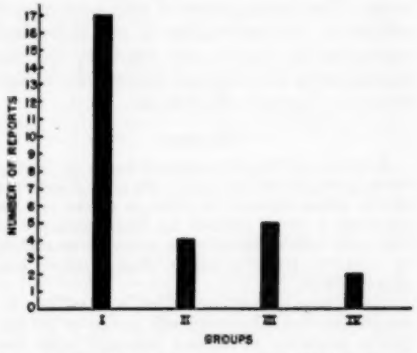

FIG. 1. THE RELATION BETWEEN THE FREQUENCY OF BAD CONDUCT REPORTS IN DIFFERENT GROUPS SEGREGATED IN TERMS OF ADJUSTMENT RATING. GROUP I HAS THE LOWEST ADJUSTMENT RATING WHEREAS GROUP IV HAS THE HIGHEST RATING.

[8] Copies of this test can be obtained on the basis of special request to the Bureau of Industrial Psychology, Extension Division, University of Wisconsin.

successive groups are not significant, but when Groups I and IV are compared, using chi-square technique, a difference significant at the .01 level was obtained. Similarly, the difference between Groups I and II combined and Groups III and IV combined is significant at the .05 level of confidence.

Inasmuch as conduct factors were determined mainly by guards and the rating by work supervisors, the present aspect of the study suggests certain uniformities of determination of adjustment in prison by those who observe the behavior of the inmate.

The main effort in this study involved determination of the correspondence between measured personality characteristics and the rated adjustment of the inmate. Figure 2 summarizes the mean scores on the subscales

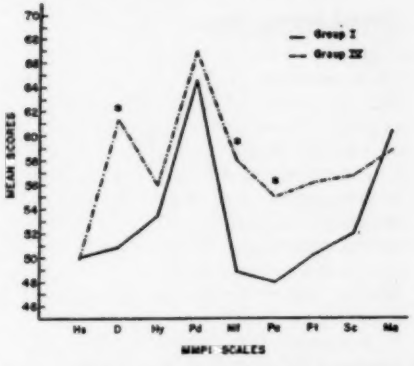

FIG. 2. THE DIFFERENCES IN PRISON INMATES SHOWING HIGH RATED ADJUSTMENT IN PRISON (BROKEN LINE) AND POOR RATED ADJUSTMENT (SOLID LINE) ON THE SCALES OF THE MINNESOTA MULTIPHASIC PERSONALITY INVENTORY.

of the MMPI for inmates of Group I and Group IV as defined above. In terms of these two groups, the inmates judged to be more maladjusted in prison (Group I) tend to give generally lower scores on most of the subscales of the inventory. It will be noted that the inmates of Group I give scores below those of Group IV on all scales of the inventory except for the *Hs* and *Ma* scales. On three of the subscales, the differences are at or beyond the .05 level of confidence when these differences are tested in terms of a *t* test of significance. The scales giving such differences are the Depression scale (.01 level), the Masculinity-Femininity scale (.05 level) and the Paranoia scale (.03 level). In terms of these comparisons, therefore, inmates judged as showing poor adjustment in prison appear more "normal" on the MMPI.

Further observations carried out indicated that inmates rated as more maladjusted were significantly younger. This group of men also contained a significantly larger number of single men.

DISCUSSION AND CONCLUSION

A special type of rating scale has been designed for study of the adjustment behavior of prison inmates. The scale designed, called the Adjustment Analysis Rating Scale, has been observed to provide a practical means of obtaining relatively consistent judgment of adaptation of the inmate to prison life. Ratings obtained from supervisors with the scale correspond in significant ways to extreme incidences of conduct as observed by guards in prison.

When the rated adjustment of the prison inmate is compared with personality inventory scores obtained shortly after incarceration, it is observed that individuals judged to be most maladjusted in prison display, in comparison to the better adjusted prisoners, significantly lower scores on the Depression, Masculinity-Femininity, and Paranoid scales of the Minnesota Multiphasic Personality Inventory. On all scales of this inventory except the Hypochondriasis and Hypomania scales, the inmates judged as maladjusted secure lower scores than those judged as being best adjusted in prison. Some general confirmation for these discrepancies in test-measured adjustment and judged adjustment in prison is to be found in an associated investigation of Gallenbeck (1), who observed evidences of the adverse effect of prison life on general personality adjustment as measured by the MMPI. In this relation, Horsch and Davis (2) report positive correlations between misconduct in prison and all "positive" traits (except sociability) on the Bernreuter Personality Inventory, and Walker (4), using the Terman-Miles "M–F" test, found consistent rule violators in prison to be more masculine than those who conform while incarcerated.

These investigations point up the fact that

the special environment of the prison may produce demands on the behavior of the individual which favor modes of conduct objectively appraised as essentially maladaptive in character outside of prison. The implications of this circumstance of institutional adjustment to prison are serious in terms of the newer concepts of rehabilitation as applied to the convicted criminal. These results suggest that the younger inmate in order to meet the requirements of prison life is required to learn modes of behavior which are essentially nonadjustive in character. The impact of this learning on post prison adjustment of the inmate is presumably to increase the difficulties already inherent in that status.

Although the present study has not been designed to be comprehensive for prison work, the findings suggest the necessity for various types of independent measures in the prediction and specification of individual adjustment of the prison inmate. It also suggests that findings in other areas of behavioral pathology could be profitably employed in penology.

REFERENCES

1. GALLENBECK, C. G. The effects of prison confinement upon personality adjustment of inmates of Waupun State Prison. Unpublished master's thesis, Univer. of Wisconsin, 1948.
2. HORSCH, A. C., & DAVIS, R. A. Personality traits and conduct of institutionalized delinquents. *J. Crim. Law Criminol.*, 1938, **29**, 241–244.
3. MCKINLEY, J. G., & HATHAWAY, S. R. The identification and measurement of the psychoneuroses in medical practice: "The Minnesota Multiphasic Personality Inventory," *J. Amer. Med. Ass.*, 1943, **122**, 161–167.
4. WALKER, E. L. The Terman-Miles "M-F" test and the prison classification program. *J. genet. Psychol.*, 1941, **59**, 27–40.

Received August 9, 1951.

STABILITY OF THE SELF-CONCEPT AS A DIMENSION OF PERSONALITY [1]

JOHN J. BROWNFAIN

Veterans Administration Regional Office, Detroit, Michigan [2]

WHEN an individual is assigned the task of evaluating himself, whatever the method of this evaluation, he inevitably makes reference to a system of central meanings that he has about himself and his relations to the world about him which we call the self-concept. Every evaluative statement a person makes about himself may be thought of as a sample of his self-concept, from which may be inferred certain properties of that self-concept. This study is concerned with one property of the self-concept—its stability.

The individual, as James (7) pointed out many years ago, has many selves. The individual might, for example, conceive of the self that he really believes he is, the self he realistically aspires to be, the self which he believes is perceived by others, the self he hopes he is now, and the self he fears he is now. The self-concept is a configuration of these and of other possible self-definitions and the stability of the self-concept derives from interrelations among these various ways of defining the self. In this study, stability of the self-concept is measured in terms of the discrepancy between two definitions of the self: the self as it is "positively" conceived and the self as it is "negatively" conceived.

The relation between stability of the self-concept and adjustment has been spoken of in a variety of ways. While the two quotations which follow do not contain the word stability or the term self-concept, what they generally convey is highly relevant to the concept of stability which will be explored in this study.

Cameron remarks: "The basis of much frustration and many conflicts is in this universal circumstance, that no man ever fuses all his self-reactions together into a single, unambiguous, coherent whole" (4, p. 102).

Rogers states: "It would appear that when all of the ways in which the individual perceives himself—all perceptions of the qualities, abilities, impulses, and attitudes of the person, and all perceptions of himself in relation to others—are accepted into the organized conscious concept of the self, then this achievement is accompanied by feelings of comfort and freedom from tension which are experienced as psychological adjustment" (9, p. 364).

One need not go far afield in the literature of psychology, psychiatry, or psychoanalysis to find concepts that would be relevant to our concept of stability and its relation to adjustment. Terms such as "integration," "consistency," "differentiation," "style of life," "wholeness of personality" are all suggestive of the same underlying variable which we call stability of the self-concept.

METHOD AND DESIGN

Subjects

The subjects (*Ss*) of this research were 62 members of two men's cooperative houses at the University of Michigan. This group, compared to the student population at the University of Michigan, might be characterized as of lower than average socioeconomic background, of liberal, democratic outlook, of diverse ethnic origin, and of above average academic achievement. The *Ss* ranged in age from 19 to 31, with a mean age of 23, and in academic level from sophomore to graduate. All *Ss* had a minimum of five months' membership in the cooperative and were, therefore, considered to be sufficiently acquainted for purposes of mutual evaluation.

Measuring the Self-Concept

The instrument which was devised as the measure of the self-concept in this study will be referred to as the Self-Rating Inventory. This inventory consists of the following 25 items, each of which is

[1] This report is based upon a dissertation submitted to the University of Michigan in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The writer is indebted to Dr. E. Lowell Kelly, who directed the dissertation, for his generous advice and criticism, and to the men of Michigan House and Owen House at the University of Michigan for fulfilling with such zeal the role of subjects.

[2] Reviewed in the Veterans Administration and published with the approval of the Chief Medical Director. The statements and conclusions published do not necessarily reflect the opinion or policy of the Veterans administration. They are the results of the author's own study.

defined at the high and low end by a brief paragraph on the printed form:

1. Intelligence
2. Emotional Maturity
3. General Culture
4. Social Poise
5. Physical Attractiveness
6. Neatness
7. Sociability
8. Generosity
9. Manners
10. Cheerfulness
11. Consistency
12. Sincerity
13. Initiative
14. Trustfulness
15. Flexibility
16. Sportsmanship
17. Individuality
18. Interest in Opposite Sex
19. Self-Understanding
20. Dependability
21. Understanding of Others
22. Self-Acceptance
23. Popularity
24. Prestige
25. Over-all Adjustment

In developing an instrument which was to serve as a measure of the self-concept an effort was made to sample as widely as possible the significant attributes of self-regard. Many of the items of the Self-Rating Inventory were adapted from the scale developed by E. Lowell Kelly and his associates (8) in their research on the selection of clinical psychologists—a scale, which, in turn, was partly based upon the factors of personality developed by Cattell (5). In addition, a number of variables were added to the Self-Rating Inventory because, according to pretest, they seemed pertinent to the requirements of this study.

Using the cooperative house membership as a reference group, each *S* rated himself four successive times on this inventory, in four different frames of judgment, each one designed to yield a different type of self-concept as follows:

1. The "private self": The most accurate estimate of himself as he really believed it to be.

2. The "positive self": A rating of himself slanted positively. Here the subject gave himself the benefit of every reasonable doubt while still conceiving of this as a believable self-picture. This is the self as he really hoped it was.

3. The "negative self": A rating of himself slanted negatively. Here the subject denied himself the benefit of every reasonable doubt while still conceiving of this as a believable picture of himself. This is the self as he feared it really was.

4. The "social self": The most accurate estimate of himself as he believed other people in the group saw him.

Ratings were made on a scale of from 1 to 8 in the private and social frames of judgment. The high end of the scale was expanded to 9 for the positively slanted ratings and the low end to 0 for the negatively slanted ratings to encourage shift in those individuals who had already given themselves extreme ratings in the private self-ratings.

The *Ss* made each set of ratings independently of the others. As a function of the unreliability of the scale, an individual could rate himself on a particular item higher in the negative and lower in the positive than in the private frame of judgment. On the few occasions that such errors occurred (the frequency being 4.6 per cent of the possible opportunities), the negative or positive rating was corrected to equal the private self-rating.

The odd-even reliability of the inventory, using ratings in the private frame of judgment, is .90 (estimated by the Spearman-Brown formula) for the sample of 62. An individual who sees himself unfavorably on some of the items is very likely to see himself unfavorably on the others. Rating behavior on this inventory appears to be essentially a reflection of the general level of self-esteem, the degree to which the individual accepts and values himself.

The Measure of Stability

Stability of the self-concept is operationally defined as the differences between positive and negative self-ratings on each item summed over all the items on the inventory without regard to sign. This difference score will be referred to as the "stability index." The larger this discrepancy score the more unstable the self-concept is assumed to be. The *Ss* with the largest discrepancies between ratings on the positive and negative self-ratings will be referred to as "the unstables"; those with the lowest discrepancy scores will be referred to as "the stables."

Theoretical considerations led us to derive a second discrepancy score from self-ratings in two different frames of judgment. It is operationally defined as the difference between private and social self-ratings on each item summed over all the inventory items without regard to sign. The greater this difference, the more discrepant is the individual's private concept of himself from the concept he believes others to have of him. We assume that when this difference is extreme, the individual cannot "be himself" in social situations, but only a facade of himself. It seems reasonable to assume that whether he believes the group overvalues him (which means that he privately undervalues himself) or undervalues him, his interpersonal relations will be fraught with possibilities of conflict and maladjustment. This discrepancy score is accordingly called the "social conflict" index.

Hypothesis

According to our general hypothesis, we expect to find that subjects whose self-concept is comparatively stable (i.e., whose negative and positive self-ratings are minimally discrepant) are individuals who, according to several criteria evaluative of adjustment, are the happier, more adequate members of the group. We expect to find, on the other hand, that individuals with an unstable self-concept will be the more poorly adjusted, unhappier members of the group.

Stability, unfortunately, cannot be examined without due regard to its pathological kinsman, rigidity. May not, after all, stability be rigidity disguised? Our hypothesis assumes that stability of the self-concept reflects an integrative function rather than rigidity of personality. To justify this assumption it was necessary to identify and then to eliminate rigid *Ss* whose stability was actually a pseudo-stability reflecting their intolerance of ambiguity

about the self. The instrument which seemed appropriate for this purpose was the F (Predisposition for Fascism) Scale developed by Frenkel-Brunswik and her associates (2) in their researches on the authoritarian personality. According to Frenkel-Brunswik, high F people, those with marked authoritarian trends in personality, are also people intolerant of ambiguity about themselves and by that criterion people who might be described as rigid.

The Pearson correlation between F Scale scores and the stability index for the sample of 62 is $-.25$ ($p<.05$). In accordance with theory, the lower (the more stable) the stability index, the higher the F Scale score. The 15 cases constituting the top quarter of the F Scale distribution were considered to be sufficiently rigid, or more accurately, sufficiently under the suspicion of rigidity, to qualify for elimination prior to testing the hypothesis. Of the 15 *Ss* thus characterized, 12 fell into the stable half of the stability index distribution, and three fell into the unstable half.

TABLE 1

PROPERTIES OF DISTRIBUTIONS OF STABILITY INDEX AND F SCALE SCORES FOR THE STABLE AND UNSTABLE GROUPS AND FOR THE TOTAL SAMPLE

GROUP	*N*	MEAN OF STABILITY INDEX DIST.	RANGE OF STABILITY INDEX DIST.	*SD* OF STABILITY INDEX DIST.	MEAN OF F SCALE SCORES	RANGE OF F SCALE SCORES
Stable	15	37	17 to 48	8	72	37 to 102
Unstable	15	86	71 to 124	14	69	46 to 96
Total	62	58	14 to 124	22	85	37 to 147

Selection and Composition of Stable and Unstable Groups

Following the elimination of the 15 cases identified as rigid from the sample of 62, the remaining 47 cases were ordered according to the size of the stability index. The upper and lower approximate thirds of these 47 cases (i.e., the 15 cases with the lowest and the 15 with the highest stability scores) were then selected. It is in terms of the differences between these two groups on various criteria of adjustment that the hypothesis will be tested. The characteristics of the stable, unstable, and total groups according to the stability index and the F Scale are presented in Table 1.

Since the stable and unstable groups were selected from extremes of the stability index distribution there is naturally no overlap between these groups on the stability index. The F Scale scores for the stable and unstable groups are virtually equivalent; presumably these two groups are equally free of rigidity as defined by the F Scale.

Demographic data collected on these groups show that there is essentially no difference between the stable and unstable groups in the following respects: age (mean for both groups, 23 years), number of veterans (10 in each group), and number of *Ss* referring to the co-op as a group of importance to them (9 in each group). The mean period of membership in the cooperative is also identical for both groups (14 months), with a range for the stables of 5 to 40 months and for the unstables of 5 to 50. Thus, both groups had equal opportunity to know other members of the cooperative and in turn to be known by them.

Criteria of Adjustment

There are four sources of data which provide measures relating to adjustment. These are:

1. Self-ratings in the four different frames of judgment;
2. Subject's scores on the Guilford-Martin Inventory of Factors GAMIN (6);
3. Group evaluation of the subject on ten variables;
4. Subject's predictions of the evaluation;
5. Subject's evaluation of other members of the group.

An individual's adjustment may be said to reflect how well he gets along with himself and his social environment. Adjustment defined in this way may be measured most directly by sources 2 and 3 above. The GAMIN tells us something about how well the individual gets along with himself while the group evaluation tells us something about how well he gets along in his social environment. The group evaluation yields the most objective criterion of adjustment since it is a source external to the *S*'s self-evaluation. The GAMIN, while not immune from the bias inherent in self-evaluation, is nevertheless a standardized instrument developed to measure certain aspects of adjustment.

In contrast, sources 1, 4, and 5, dealing with *S*'s behavior in evaluating himself and others, provide what might be properly considered secondary measures of adjustment. They do not measure adjustment directly, but they do permit us to make inferences about adjustment.

Group Evaluation Procedures and the Measures Derived from Them

Group members evaluated each other on the following ten variables:

1. Knowing (How well do you know him?)
2. Liking (How well do you like him?)

3. Prestige (How much prestige does he have in your eyes?)
4. Popularity (How popular do you think he is with the group?)
5. Consistency
6. Intelligence
7. Flexibility
8. Understanding of Others
9. Dependability
10. Over-all Adjustment

These variables, except for 1 and 2, are also among the items of the Self-Rating Inventory and are defined in approximately the same way.

In order that each *S* would have a common standard for evaluating others, it was decided to use a ranking procedure rather than free ratings. It was felt that if each *S* ranked 16 randomly assigned group members and was in turn evaluated by 16 group members, an adequate measure of the group consensus would be obtained. Each of the 16 rankings made on a given individual was in terms of a differently constituted sample so that, ultimately, each individual was compared and then ranked in relation to the entire house membership. To reduce the strain of making fine discriminations, and also to establish a numerical frame of reference consistent with the eight-step scale used in taking the Self-Rating Inventory afterwards, *Ss* were instructed to assign two individuals to each rank from 1 to 8.

On the same forms the *Ss* used to evaluate group members, they also recorded predictions of the average ranking that would be assigned to them by the entire membership of the co-op house. These predictions took three forms: most accurate estimate, highest estimate, and lowest estimate of the average ranking on each variable. It is possible to obtain from these data a measure similar to the stability index: the difference between the highest and lowest estimate, summed over all ten variables without regard to sign. The resulting index was identified as "stability-of-prediction." It constitutes a measure of the stability of another aspect of the self, the social self, the self as one believes it is perceived by others.

We were also interested in investigating the individual's ability to predict how the group saw him. The difference between an *S*'s most accurate estimate of group rankings of him and the actual mean of rankings made by the group members (summed over all ten variables without regard to sign) provides one measure of the *S*'s social insight. This measure will be referred to as "insight I."

Eight of the variables on which *Ss* ranked each other were also items in the Self-Rating Inventory. On these eight variables a second measure of insight was obtained: the difference between self-ratings in the private frame of judgment and the mean of group rankings (summed over all 8 variables without regard to sign). This measure is an index of the congruence between the self as it is privately seen and the self as others perceive it. This discrepancy score will be referred to as "insight II." This measure is operationally similar to the way Sears (10) defined insight in his study of projection.

As would be expected, both indices of insight tend to measure the same psychological variable, as shown by a correlation of .63 ($N=62$; $p<.001$).

Some Characteristics of the Stability Index

Since the stability index is a measure unique to this study, it is important to know something about some of its characteristics, particularly its reliability and its susceptibility to possible artifacts. In this section, these characteristics of the stability index will be reported in terms of the total sample. In the section dealing with the test of the hypothesis in terms of differences between the stable and unstable groups, further light will be cast on the stability index by showing its relations to various measures of adjustment.

The odd-even reliability of the stability index, estimated by the Spearman-Brown formula, is .93 for the total sample of 62. Such a reliability coefficient indicates that, in spite of the diversity of items making up the Self-Rating Inventory, stability of the self-concept affects self-rating behavior in a general way over all the inventory.

The tests of the hypothesis will deal in some detail with the relation between the stability index and private self-ratings. For the entire sample of 62, the correlation between the two measures is $-.25$ ($p<.05$). Lower self-esteem is more likely to be associated with an unstable self-concept.

The question immediately arises: might not the stability index be in part an artifact of the level of self-rating? That is, as the level of private self-rating rises, there is less room to shift upward in the positive frame of judgment; hence the stability index is apt to be smaller. To test this, the mean of self-ratings was computed for each of the 25 *items* (averaging across the 62 *Ss*). These means were then correlated with the means of the stability scores computed for the 25 items in a similar way. The correlation between the two is .19. Though this correlation is not significant, its positive direction provides an emphatic denial to the question of artifact. A high self-rating on a particular item does not per se reduce the stability score on that item; it is even possible for the reverse to occur.

RESULTS

While the study was still in design stage, a series of specific predictions was made relative to differences between the stable and unstable groups on the various measures of adjustment. The findings are confined for the most part to tests of those specific predictions.

The statistical technique used to test significance of differences between means of the stable and unstable groups on the criteria of adjustment is Fisher's *t*, using the appropriate small sample formula. In testing the significance of *t* values for findings, one-half of the probability curve was used since in such instances direction of difference between the stable and unstable groups was predicted.

Probabilities which attain the .05 level will be considered significant. Degrees of freedom in *t* tests is 28 unless otherwise indicated.

Self-Ratings

Table 2 shows that the stable *Ss* have both a higher level of self-rating in the private frame of judgment and a lower intertrait variability in these ratings. The stable group may be characterized, therefore, as having a greater amount of self-esteem. The stable *Ss* also see themselves in a more constant way while, by contrast, the unstable *Ss* see themselves in a more variable way from trait to trait, suggesting that they have an ambivalent attitude toward themselves which is expressed in a lower rating of certain attributes of the self and what might be a compensatory accentuation of other attributes of the self.

On 21 of the 25 items of the Self-Rating Inventory, stable *Ss* rate themselves higher than do unstable *Ss* (data not shown). Differences on five of these items attain significance ($p<.05$). These are Cheerfulness, Self-Acceptance, Over-all Adjustment, Consistency, and Trustfulness. On none of the four items on which the unstables rate themselves higher is the difference significant. Of the four items (Self-Understanding, Intelligence, General Culture, and Dependability), it is the difference on Intelligence that comes closest to being significant. Since the grade point average of the unstable *Ss* approximates a B and that of the stable *Ss* a C+ (difference significant at .01 level), it follows that there is some reality basis for the higher self-evaluation of the unstable *Ss* on Intelligence.

The reader will recall that the difference between the way *S* sees himself (private self) and the way he believes others see him (social self) was labeled the social-conflict index. As Table 2 indicates, *Ss* in the unstable group, as compared to the stable group, show an enormous incongruity between the private perception of the self and the perception of the self they attribute to other people. According to the rationale of the social conflict index, the greater this discrepancy score the larger the "breeding ground" for difficulty in interpersonal relations. Later evidence will show that the unstable *Ss* are indeed less well adjusted socially than the stable *Ss*.

Guilford-Martin Inventory of Factors GAMIN

Except for the GAMIN, all measures of adjustment were derived from instruments especially developed for purposes of the study. The findings on the GAMIN are therefore particularly noteworthy. While it was felt that the stable group would in general have "healthier" GAMIN profiles, specific predictions were made in their favor only with respect to the I and N factor scores.

The evidence from the GAMIN, as Table 3 shows, altogether substantiates the fact that the stable *Ss* are better adjusted than the unstable *Ss*. On only one factor (G) does the difference favor the unstable *Ss* and this difference is not significant.

On the other four factors, the stable *Ss* have the "healthier" scores. They score somewhat more ascendant and more masculine. They are less afflicted by inferiority feelings and by nervousness. The significant difference on the N factor is especially noteworthy since it was established that this factor is not significantly correlated with the mean of private self-ratings as are the A and I factors.

In scoring the GAMIN, a technique de-

TABLE 2

DIFFERENCES BETWEEN STABLE AND UNSTABLE GROUP ON LEVEL AND VARIABILITY OF PRIVATE SELF-RATINGS AND ON THE SOCIAL CONFLICT INDEX

	MEAN		Diff.	*t* VALUE	$p<$
	STABLE	UNSTABLE			
Mean of Self-Ratings	6.09	5.64	.45	1.802*	.05
SD of Self-Ratings	1.25	1.62	−.37	3.064*	.005
Social Conflict Index	18.58	34.95	−16.37	4.052*	.001

* Probabilities for these *t* values are based upon one-half of the probability curve since direction of difference was predicted before analysis of data.

TABLE 3

DIFFERENCE BETWEEN STABLE AND UNSTABLE GROUPS ON THE INVENTORY OF FACTORS GAMIN *

	MEAN				
	STABLE	UNSTABLE	Diff.	*t* VALUE	$p<$
C-scores** on General Activity	4.07	4.60	−.53	.882	.40
" " Ascendance	5.50	4.60	.90	1.304	.30
" " Masculinity	7.36	6.13	1.23	2.014	.10
" " Inferiority (lack)	4.93	3.47	1.46	2.028***	.05
" " Nervousness (lack)	5.29	3.73	1.56	2.108***	.025

* Since one member of the stable group did not complete the GAMIN, $df=27$.

** The median C-score for the 160 university students on which the Guilford-Martin norms are based is 5.

*** Probabilities for these *t* values are based upon one-half of the probability curve since direction of difference was predicted before analysis of data.

veloped by Ernest Tupes[8] to measure the amount of "facade" (i.e., the tendency of a subject to "fake" responses in order to present himself in a more favorable light) was used. The facade scores for the stable and unstable groups as well as for the entire sample are so low that, to quote Tupes (private communication) the subjects must be considered to have approached the GAMIN with "almost painful honesty." Such a finding enhances the validity of the GAMIN scores which were obtained and suggests furthermore that the *Ss* were very likely optimally sincere and cooperative in completing other instruments in accordance with instructions.

The Group's Evaluation of Each Subject

So intimate is the relationship between self and the social environment that it may be asserted that there is no maladjustment in personality that is not somehow reflected in maladjustment in interpersonal relations. One would therefore look to the group's evaluation of each individual as a crucial criterion of his adjustment.

We shall first examine the differences between mean ranks of the stable and unstable *Ss* on the 10 group evaluation variables as presented in Table 4. Nine of the 10 differences favor the stable group. Some of these differences may result from the intercorrelations among these variables due to "halo." Even so, it is interesting that the three significant differences are on variables most pertinent to self-group relationships, variables measuring how well known and how well liked the individual is, and how much popularity is attributed to him.

The difference on *knowing* is especially revealing since mean length of membership

[8] TUPES, E. C. The detection of "faking" on the Guilford-Martin Personality Inventory GAMIN. Unpublished study.

TABLE 4

DIFFERENCES BETWEEN MEAN RANKS OF STABLE AND UNSTABLE GROUPS ON THE TEN GROUP EVALUATION VARIABLES

	MEAN		Diff.	*t* VALUE	$p<$
	STABLE	UNSTABLE			
1. Knowing (how well known)	5.50	4.73	.77	1.944*	.05
2. Liking (how well liked)	4.87	4.03	.84	1.809*	.05
3. Prestige (how prestigeful)	4.82	4.05	.77	1.645*	.10
4. Popularity (attributed to group)	5.43	4.47	.96	1.844*	.05
5. Consistency	4.55	3.86	.69	1.647*	.10
6. Flexibility	4.63	4.14	.49	1.111	.30
7. Intelligence	4.33	4.89	−.56	1.090	.30
8. Understanding of others	4.86	4.13	.73	1.460*	.10
9. Dependability	4.47	4.10	.37	.678	.60
10. Over-all adjustment	4.81	4.05	.76	1.652*	.10
Mean of all ten ranks	4.83	4.25	.58	1.563*	.10

* Probabilities for these *t* values are based upon one-half of the probability curve since direction of difference was predicted before analysis of data.

for stable and unstable groups was exactly the same.[4] The stables must be the more active socializers to be known better. It would appear that the unstables tend to withdraw from group contact. On only one out of the 10 variables are the unstable *Ss* favored and that is on *intelligence*. This is understandable in view of their superior academic achievement.

Subjects' Predictions of the Group Evaluation

We are concerned with three aspects of the predictions made by the *Ss* of the group evaluation: their level, their stability, and their accuracy. These findings are presented in Table 5 and are limited to an analysis of over-all scores rather than scores on each of the 10 group evaluation variables.

The level of predictions was based upon the "most accurate estimate" made by *S* of the group's evaluation. The *Ss* also recorded their highest and lowest estimates of the group's evaluation. From these data was derived a measure similar to the stability index called the stability-of-predictions. On all 10 variables (not shown) the stable *Ss* make more stable predictions than do the unstable *Ss*. When this behavior is considered for the total of the 10 variables, the difference between the two groups becomes highly significant. The stable *Ss* are more confident that they know what the group thinks about them.

We now deal with the third aspect of *Ss*' predictions—their accuracy. Two indices of insight into social reality have been described.

TABLE 5

DIFFERENCES BETWEEN STABLE AND UNSTABLE GROUPS ON LEVEL, STABILITY AND ACCURACY OF PREDICTION OF THE GROUP EVALUATION

	MEAN		Diff.	*t* VALUE *	$p <$
	STABLE	UNSTABLE			
Mean of ten predictions	5.25	4.86	.39	1.639	.10
Stability-of-prediction index (total for ten variables)	24.53	32.2[illegible]	−7.74	2.847	.005
Insight I index (total discrepancy for ten variables)	10.73	12.2[illegible]	−1.47	.624	.30
Insight II index (total discrepancy for each eight variables)	12.07	14.87	−2.80	1.418	.10

* Probabilities for these *t* values are based upon one-half of the probability curve since direction of difference was predicted before analysis of data.

The stable *Ss* show by way of their higher level predictions that they expect the group to evaluate them more favorably. On nine of the 10 variables (the exception being *intelligence*) the stable *Ss* make higher level predictions (data not shown). While it is only on *over-all adjustment* that the difference is statistically significant, the trend is nevertheless unmistakable. Of course, the stable *Ss* see themselves privately in a more favorable light than do the unstable *Ss* and they are in reality more highly regarded by the group as a whole. Their predictions undoubtedly reflect both their self-esteem and the reality picture.

[4] Subjects were asked to report how many people they knew well or fairly well among the 40 to 45 members belonging to their co-op house. The difference between the stables and unstables in this regard is striking. The stables report knowing on the average 31 people in the co-op, the unstables 21 (*t* value 2.472; $p < .01$).

Table 5 shows for both Insight I and Insight II a slight tendency for the stable *Ss* to estimate more accurately the group consensus about them.

Looking at Insight II, we see that the stable *Ss* tend to see themselves privately in a way which is more congruent with the social estimate of them than is true of the unstable *Ss*. Even though the difference for the small *N* fails to attain significance, it is interesting to speculate about the possible meaning of this. It appears that what the stable *Ss* really think about themselves comes closer to what the group really thinks about them. When this finding is related to the fact that the stable *Ss* are also more apt to see themselves in a way congruent with the way they think other people see them (as discussed in the analysis of the social conflict index), we may infer that the stable *Ss* are

more free to "be themselves" in interpersonal situations and that others are better able to evaluate them as they really are.

Subjects' Evaluations of Others

The group evaluation data not only tell us what the group thinks about the individual, but they also permit us to draw certain inferences about the way the *S* perceives the people he evaluates. The reader will recall that each *S* ranked 16 members of the cooperative on ten different variables. For each *S* separately a rank-order correlation (rho) may be computed between any two sets of rankings. For example, a rho may be computed between the way a given *S* ranks his 16 people on how well he knows them, and the way he ranks them on how well he likes them. Rho's for all the stable *S*s and the unstable *S*s respectively may be averaged (following a transformation of rho's into normally distributed variates) to find out whether the stable and unstable groups differ in the degree to which they associate any two sets of rankings. Differences between the stable and unstable groups on three pairs of rankings treated this way are presented in Table 6.

The unstable *S*s, as shown by their lower mean rho between *liking* and *knowing*, are less likely to like the people they know well (or, conversely, to know well the people they like best) than are the stable *S*s. We may speculate that the maladjusted person is more likely to take out his conflicts upon those he knows well and becomes emotionally involved with, a state of affairs which leads to alienation. He may end up preferring individuals whom he knows less well and against whom he is protected by social distance.

The unstable *S*s are less likely to like people they consider popular in the group than is true of the stable *S*s, as shown by their lower mean rho between *liking* and *popularity*. Perhaps envy, values rooted in rationalizations (e.g., popular people are superficial), or a sense of painful contrast could account for this.

The relationship between *liking* and *prestige* was investigated because it was felt that less secure people (the unstable ones) would feel threatened by or envious of prestigeful figures and would be less likely to like them than in the case of the more secure (stable) people. The difference, although in the predicted direction, is not significant.

To summarize, *S*s with the more stable self-concepts are better adjusted than those with less stable self-concepts on all of our criteria of adjustment whether these criteria are derived from self-evaluation on the Self Rating Inventory and the Guilford-Martin Inventory of Factors GAMIN, from the group's evaluation of the *S*, or from the *S*'s evaluation of group members. Certain differences do not attain statistical significance. Yet, in spite of the small number of *S*s involved in testing the hypothesis, all the differences are in the direction of the theoretical prediction. It is this consistency of results obtained from an extensive analysis which provides the most impressive basis for accepting the hypothesis concerning differences between the stable and unstable groups in our co-op sample.

TABLE 6

DIFFERENCES BETWEEN STABLE AND UNSTABLE GROUPS ON RANK ORDER CORRELATIONS BETWEEN PAIRS OF GROUP EVALUATION RANKINGS

	Mean		Diff.	*t* Value *	$p <$
	Stable	Unstable			
Liking and knowing	.62	.44	.18	1.755	.05
Liking and popularity	.62	.47	.15	1.569	.10
Liking and prestige	.72	.69	.03	.411	.35

* Probabilities for these *t* values are based upon one-half of the probability curve since direction of difference was predicted before analysis of data.

Discussion

The hypothesis was tested in terms of criteria of adjustment. Some of the main ingredients of this "adjustment" are those referring to self-esteem. Measures dealing with self-evaluation, whether the Self Rating

Inventory or the Guilford-Martin GAMIN, are more than anything else reflective of such qualities as self-adequacy, self-acceptance, self-confidence—reflective, in a word, of self-esteem. A major correlate of the stable self-concept (provided that this stability does not function in the interest of a defensive rigidity) is a high level of self-esteem. The individual with a stable self-concept is the individual who accepts himself, who values himself highly, who feels secure about himself. When he asks of himself the question "Who am I?", he is able to answer himself in a confident manner and he is generally pleased with his answer. But the individual with an unstable self-concept is deficient in self-esteem. He is uncertain about who he is and is not likely to be pleased with the answer he gives himself. It is true that such an individual may be capable of seeing himself in a highly flattering way, but this is just one pole of his ambivalence. When he thinks he's good he thinks he's very good, but when he thinks he's bad, he thinks he's horrid.

While self-esteem may be considered a correlate of stability of the self-concept, one cannot say whether it is the cause or effect. It would be more reasonable to posit a circular-causal relationship between the two. Two speculative explanations of this circular-casual relationship are suggested.

The first type of explanation posits a compensatory reaction to low self-esteem. The individual who is deficient in self-esteem must defend himself against the insecurity which accompanies this condition so that, for good defensive reasons, he is also capable of seeing himself in a positive way. People with unstable self-concepts are unable to consolidate their various self-concepts into a stable organization because they cannot tolerate the negative elements which are inherent in low self-esteem. So long as they are casting about for an acceptable self-picture, the self-concept must remain fluid.

The relationship between self-esteem and stability of the self-concept might be examined in another way. When a person evaluates himself in the positive and negative frames of judgment, his behavior is likely to be governed by some awareness of his having been at some time or other particularly adequate or inadequate in various situations. The individual lacking in self-esteem is more likely to be "situation dominated." Uncertain about what to expect from the environment, his behavior becomes more dependent upon what he perceives as its demands. If the situation is favorable enough, his self-esteem is inordinately heightened, but if the situation is unfavorable, self-esteem is inordinately lowered. Whether he is his adequate or inadequate self seems to hinge upon the eddies in the emotional environment. On the other hand, the individual with high self-esteem is less likely to see himself as dominated by a situation. His self-concept is relatively stable because he is sufficiently the master of his environment that a radical restructuring of the self-concept need not be the consequence of changing situations.

Just as evidence from self-evaluation shows that the people who have unstable self-concepts also have lower self-esteem, the findings based on the group evaluation show that the group esteems them less too. The group is likely to share the individual's attitude toward himself. This may be, it is true, partly a function of the reality about the individual—that is, he stands low on a number of traits and both he and the group may be aware of this. But on the other hand, his own attitude toward himself may generate a similar attitude in the group toward him. To illustrate this, the unstables do not, to begin with, value themselves highly, but they see the group as valuing them far less than they do themselves. This makes them fearful of the group and this fear motivates them to withdraw from the group. As a consequence they know fewer people in the group and are less well known by the group. Of course, it is not only their withdrawal from social contact which fosters the rejection. If the individual perceives the social environment as threatening, then he may respond to it with hostility which provokes counter-hostility in the group, leading to still further alienation.

The behavior of the individual with an unstable self-concept, which has as one correlate lowered self-esteem, is quite in accordance with Adlerian theory. Adler (1, p. 23) has remarked that people who feel inferior tend to withdraw from social participation until only a small group is left over for the "maneuvers aiming at the various types of superiority to expend themselves upon." There is, of course, evidence that the unstables do withdraw and that they do use the

defense of compensation which Allport (3) has called the "most interesting of all the shy handmaidens of self-esteem." However, the writer does not wish to impose the Adlerian conception on the findings of this study. There are undoubtedly other correlates of the unstable self-concept which were not, within the confines of this study, brought to light as was self-esteem. It is possible that a study using clinical techniques such as the detailed case history and projective tests would reveal other correlates of instability of the self-concept, defense mechanisms other than what has been referred to so grossly as "compensatory."

SUMMARY

The purpose of this study was to develop an operational measure of the stability of the self-concept and to demonstrate that stability of the self-concept is a dimension of personality serviceable to the work of understanding and predicting behavior.

The primary data of this investigation consist of several series of self-ratings on 25 personality variables, obtained successively under different instructions from 62 members of two men's cooperative houses at the University of Michigan. Under one set of instructions, the *S* gave himself the benefit of any realistic doubt he had about his standing on each inventory item, thus yielding a "positive" self-concept. Under another set of instructions *S* denied himself the benefit of such doubt, thus yielding a "negative" self-concept. The difference between these positively and negatively slanted self-ratings on each item, summed over all the items of the inventory without regard to sign, was the operational measure of stability. The larger this discrepancy, the more unstable the self-concept is assumed to be. This measure of stability has an estimated reliability of .93.

The hypothesis was: *Ss* who have the most stable self-concepts are, according to a number of criteria, better adjusted than *Ss* with the least stable self-concepts.

In making this hypothesis, it was assumed that stability reflects an integrative function rather than rigidity of personality. Therefore, the 15 *Ss* identified as rigid, on the basis of high scores on the F (Predisposition for Fascism) Scale developed by Frenkel-Brunswik, were eliminated from the sample before testing the hypothesis.

Adjustment was measured in terms of inferences derived from *S*'s behavior in rating himself and others, *S*'s scores on the Guilford-Martin Inventory of Factors GAMIN, and the group's evaluation of *S*. Fisher's *t* was used to test significance of differences between the means of these measures for the 15 *Ss* with the most stable and the 15 with the least stable self-concepts.

All findings support the theoretical prediction that *Ss* with stable self-concepts are better adjusted than those with unstable self-concepts. The following are the more salient findings which favor the *Ss* with more stable self-concepts ($p < .05$):

1. They have a higher level of self-esteem as manifested by a higher mean self-rating and also by a higher self-rating on the inventory item defining self-acceptance. The intertrait variability of their self-ratings is lower, indicating that their self-esteem is generalized.

2. They are freer of inferiority feelings and nervousness as measured by the GAMIN.

3. They are better liked and considered more popular by the group.

4. They see themselves more as they believe other people see them.

5. They know more people in the group and are better known by the group, indicating more active social participation.

6. They show less evidence of compensatory behavior of a defensive kind.

REFERENCES

1. ADLER, A. *The practice and theory of individual psychology.* London: Kegan Paul, Trench & Trubner, 1925.
2. ADORNO, T. W., FRENKEL-BRUNSWIK, ELSE, LEVINSON, D. J., & SANFORD, R. N. *The authoritarian personality.* New York: Harper, 1950.
3. ALLPORT, G. W. *Personality, a psychological interpretation.* New York: Holt, 1937.
4. CAMERON, N. *The psychology of behavior disorders.* Boston: Houghton Mifflin, 1947.
5. CATTELL, R. B. Confirmation and clarification of primary personality factors. *Psychometrika,* 1947, 12, 197–220.
6. GUILFORD, J. P., & MARTIN, H. G. *The Guilford-Martin Inventory of Factors GAMIN: Manual of directions and norms.* Beverly Hills, Calif.: Sheridan Supply Co., 1945.
7. JAMES, W. *Principles of psychology,* Vol. I. New York: Holt, 1890.
8. KELLY, E. L. Research on the selection of clinical psychologists. *J. clin. Psychol.,* 1947, 3, 39–42.
9. ROGERS, C. R. The organization of personality. *Amer. Psychologist,* 1947, 2, 358–368.
10. SEARS, R. R. Experimental studies of projection: 1. Attribution of traits. *J. soc. Psychol.,* 1936, 7, 151–163.

Received September 3, 1951.

THE RORSCHACH TEST AND SUGGESTIBILITY[1]

IRA M. STEISEL
University of Washington

THE widespread use of the Rorschach test and the voluminous literature on its clinical application indicate that clinical workers have accepted the technique as a valuable instrument for the description and evaluation of personality processes in the individual. It is somewhat surprising, however, to note that this general acceptance and extensive use have been based upon only relatively few studies of the relationship between the Rorschach test measures and other measures of the same personality variables. It would appear highly desirable, therefore, if a more thoroughgoing understanding of the Rorschach (and the field of personality evaluation) is to be achieved, to augment these studies with others of a similar nature.

Investigations of this type, in which the results of one test are compared with those of another, are commonly termed "validation" studies. The problem of attempting to "validate" the Rorschach (or any test for that matter) by such a procedure is often difficult because in many instances there are almost no tests of the same psychological processes that are universally accepted as suitable criterial indices. Also, the Rorschach is presumed to measure a number of personality variables simultaneously. The best that can be done, therefore, is to study the interrelations of the Rorschach and a wide variety of other tests, with the clear understanding that at our present stage of knowledge either the Rorschach or the other tests might be chosen as the criterion. The final decision as to which shall be called the criterion depends upon such other factors as internal consistency and the kinds of behavioral laws or functional relations to which the tests lead. It should be noted, however, that if a single test such as the Rorschach should correlate highly with a variety of different tests, then its use would be justified on the basis of economy since a large number of predictions could be made through the administration of but one test.

An additional factor that makes investigation along these lines difficult is the statement made by Rorschach workers that the identification of a personality characteristic by means of the test is not a piecemeal procedure. Raw scores from parts of the test are not alone adequate. Account must be taken of the interrelations of the raw scores and the patterns formed by them. In some instances, therefore, a "validation" study would be acceptable only if such relationships were considered. As Rotter (30) has pointed out, however, the Rorschach interpreter often acts as if each unique determinant had a specific meaning. Also, the patterns that apply to an unique interpretation are rarely specified. Williams (37), being fully aware of the need to consider the interrelations of the scores, attempted to investigate the personality process of "intellectual control" through the specific Rorschach test factors that have been assumed to reflect this process. As he has indicated, this procedure does not imply a denial of the existence of complex relationships; it was employed solely for convenience of verification and simplicity.

Williams (37), combining the previous suggestions of Hertz (17) and Sargent (31), suggests that Rorschach validity studies might be classified into two broad categories: (*a*) the comparison of Rorschach findings with clinical findings and (*b*) the correlation of Rorschach findings with independent and experimentally controlled behavioral criteria.

The majority of studies in this general area are of the first type suggested by Williams. The work of Rorschach (29), Beck (1, 2, 3,

[1] This paper is a major portion of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Psychology at the State University of Iowa. The writer would like to express his indebtedness to the codirectors of this dissertation, Drs. Judson S. Brown and Arthur L. Benton, for their critical suggestions and assistance during the course of this study. Special thanks are due to Drs. Woodrow W. Morris and Bertram D. Cohen (now at Indiana University) for their stimulating discussions on problems involved in the interpretation of Rorschach's test and to Mr. Ted Hunter for the construction of a major portion of the apparatus employed in this study. A summary of this paper was read at the April, 1951, meeting of the Western Psychological Association.

25), Hertz (16, 18, 19), and Klopfer (24) may be cited as examples.[2] In such studies the usual procedure is to relate judgments based on Rorschach data (or the raw data themselves) to judgments made from other data; e.g., personality descriptions, psychiatric diagnoses, etc.

The studies encompassed in Williams' second category have utilized, for the most part, criteria that are defined independently of the Rorschach, that are measurable, clear-cut, and subject to a minimum of judgmental error. These factors, plus the fact that the technique readily allows for experimental control and manipulation, would seem to indicate that its exploitation might prove profitable in the discovery of additional lawful relationships.

The present investigation is considered to be an example of the type of study included in Williams' second category. In it, Rorschach measures hypothesized to reflect the personality variable of suggestibility are compared with other previously used measures of the same variable obtained in an experimental situation.

THE PROBLEM

Rorschach experts agree in assuming that individuals with responses determined in part or entirely by the color in the figures respond readily to affective stimulation from the external environment. Depending on the relative influence of form on the color-determined responses and other pertinent variables in the record, various descriptive phrases have been employed to specify the way in which this reactivity to the environment will be expressed. Irritability, impulsivity, egocentricity, emotional immaturity, and suggestibility are terms that have been frequently used. Rorschach indicates that, among others, the personality facet of suggestibility is indicated by the presence of responses scorable as *CF*. He states:

> Affective suggestibility is represented in this test by the *CF* responses; *M*, *FC*, and *C* answers may or may not be present; in any case there must be a predominance of *CF*'s, either absolutely or relatively. Egocentric extratensive subjects are the most suggestible, and are particularly influenced by emotional suggestions. The greater the number of *M*'s in the experience type formula, the less suggestible is the subject, but if suggestive influences are at all effective in such individuals they are more lasting. A fact already established is reaffirmed in this connection, namely, that the greater the capacity for conscious reasoning, the less suggestible is the subject, and conversely, the greater the affective lability, the greater the suggestibility of the subject. . . . Coartated types are not very suggestible (29, pp. 100–101).

More recently, Beck (3, p. 29) and Bell (4, p. 123) have indicated that one of the personality counterparts of the *CF* Rorschach response may be suggestibility.

The concept of suggestibility has been extensively studied and discussed by previous investigators.[3] Several of the more recent studies have suggested, since various tests of suggestibility may not correlate well with one another, that it is not a unitary, general trait but is composed of several distinct personality variables (7, 10, 11, 12, 13, 15, 21, 28). Coffin (11), after reviewing the literature in this area, concludes that suggestibility is a general term and that one must "specify the particular conditions under which the degree of suggestibility has been observed." He indicates, further, that most writers would accept the proposition that suggestibility is some " 'uncritical' or 'unreasoned' response . . ." and that the definitions proposed by various authors imply that the subject "uncritically" accepts a proposition or course of action foisted upon him by the experimenter.

Because of the several quite different properties that may be designated by the term, it would seem most wise in studying the relationships between Rorschach variables and suggestibility to use different measures of that variable. The following tests, the majority of which have been used by previous investigators, were employed in the present study:

1. The Postural-Sway Test. This test was designed by Hull (20) in 1929 and has been used by a number of later investigators (cf. 12, 13, 26). The apparatus permits the recording of movements of an individual who is initially told to stand still with his eyes closed (or blindfolded) and has repeated to him the suggestion that he is falling forward. The extent of the resultant movement is taken as the index of the subject's suggestibility. Eysenck (12) has indicated that this is the best single measure of what he has termed "primary suggestibility" (ideo-motor).

2. The "Ink-blot Suggestion Test." This test was described by Eysenck and Furneaux (13). In their study a Rorschach card was shown to each subject and he was presented with two "common" responses and four "inapplicable" responses with the comment that "some people think it looks like that." Each subject was asked if he could see something resembling those percepts in the figure.

[2] For extensive reviews of validity studies of both types, see Benjamin and Ebaugh (5), Hertz (17, 18), Sargent (31), Schneider (32), and Williams (37).

[3] For extensive reviews of the literature on this topic, see Bird (7), Coffin (11), Hull (21), and Murphy, Murphy, and Newcomb (28). Bird presents an extensive bibliography through the year 1939 (6).

The authors indicate that the number of such percepts accepted by the subject was his suggestibility score. They, further, indicate that this is the best index of "secondary suggestibility" (gullibility).

3. Three measures derived from two autokinetic situations were used.[4] The work of Sherif (33), Bray (10), Kelman (23), and Bovard (9) indicates that the apparent movement of a pin point of light in an otherwise totally dark room might be utilized to investigate personality processes and social psychological phenomena. They found that, in general, the average extent of movement reported by a subject, when in the situation alone, could be altered to conform to a group average or an average supplied by a planted subject with prestige in a retest situation.

Although Sherif and Bovard do not speak of the obtained measures in terms of suggestibility, they do speak of the individual as accepting social norms. The results of their studies, however, indicate that some measures obtained from the autokinetic situation might be indicative of the degree of suggestibility of an individual. The measures that were employed in the present study were: (*a*) The amount of reported movement when the experimenter suggested that the light would move.[5] (*b*) The degree of correspondence of the individual's estimates, in a retest situation, to a series of arbitrary estimates offered at that time. Kelman (23) and Bray (10) have recently used this latter as a measure of suggestibility. (*c*) The time taken by the subject to indicate the perception of movement after the light had been turned on. Although this specific measure had not been employed previously, it was believed that the more highly suggestible individuals would respond more quickly. It was assumed that the more critical individual would take a longer time in making his judgment. This hypothesized relationship is also consistent with another interpretation of the presence of *CF* responses in the Rorschach protocol, i.e., impulsivity.[6]

For this study, the Rorschach measures that were hypothesized to be related to suggestibility were derived from the previously quoted statements of Rorschach, as well as from the writings of Beck (3), Bell (4), Klopfer and Kelley (24), Williams (37), Hertz (16) and others. The following testable hypotheses were formulated: Other things being equal,

1. A positive relationship exists between the absolute number and/or per cent of responses scorable as *CF* and some measure of suggestibility. This follows directly from the statements of Rorschach, Beck, and Bell.

2. The lower the absolute number and/or per cent of responses scorable as *M*, the greater would a suggestibility score be. This, too, follows from Rorschach's statement.

3. A negative relationship exists between a measure of the amount of intellectual control displayed when an individual is confronted by emotional stimuli and some measure of suggestibility. This, it would seem, follows from the statements of various authors that suggestibility is an uncritical emotional response. Williams (37) has used the $F + \%$ on colored cards as an index of intellectual control in an emotionally charged situation and this was used as such in the present study.

4. A negative relationship exists between a measure of conscious control and some index of suggestibility. Rorschach indicates that this is the anticipated relationship. Inasmuch as Klopfer and Kelley (24) describe the $F\%$ as a measure of conscious control, this measure was employed as such in the present investigation.

5. The greater the amount of constriction reflected in the Rorschach, the less suggestible would the individual be. Rorschach specifically indicates this (see above). Since the $A\%$ has been offered as one measure of constriction (3, p. 16), this measure was so used here.

6. The lower the critical ability of an individual, the greater the amount of suggestibility. According to the standard interpretation, an individual who evidences high accuracy of form-perception displays good critical ability. This particular characteristic is, according to the previously quoted statements, low in the suggestible individual. The $F + \%$ has been described as an index of accuracy of form-perception (3) and was so employed in this study.

7. A positive relationship exists between the amount of conformity of thinking displayed in the Rorschach protocol and the degree of suggestibility (especially of the prestige type). The absolute number and per cent of responses scorable as "popular" have been used as indices of conformity of

[4] Varvel (35) proposed the use of the autokinetic situation and other experimental procedures as a means of investigating correlates of Rorschach variables.

[5] It was argued that since the light does not move, the report of movement, after such instructions are given, is an indication of the subject's suggestibility. No other writer, as far as the present author is aware, has spoken of this measure as such.

[6] Schneider (32, pp. 499–500) has recently discussed this study as an example of one using operationally defined situations to validate Rorschach variables. He argued that the criterion measures employed here are "not as universally acceptable as is the relationship between intellectual control and the situation used by Williams." It is agreed with Schneider that the relationship between the criterion measures used here and still other measures should be discovered. This has been done by previous workers, in part at least, for the majority of indices employed here. It should be noted, however, that Williams was apparently the first one to utilize his stress situation as a test of intellectual control. Although it would "seem" to be an adequate measure of this personality variable, the present author would argue that face validity is not too adequate for a scientific study of personality. If face validity were sufficient, then there would be no need for investigating behavioral correlates of Rorschach responses and both this study and a major portion of Schneider's paper would be pointless. It is here argued that there is still a need for finding the relationship between Williams' situation and other tests of intellectual control. Strictly speaking, his study has shown nothing more than that one could predict, with a certain degree of accuracy, from the specific Rorschach measures to a specific test situation and that an experimental test of Rorschach hypotheses is feasible.

thinking (3, 16, 24) and were used as such in this study.

8. The greater the amount of suggestibility of an individual, the greater will be the amount of extratension he will display in his Rorschach record. Williams (37) has introduced the average color score (the sum of the conventional weights of the various color responses divided by the absolute number of responses determined by color). This ratio weights the color responses in such a way that the higher values are indicative of less control over emotional factors and more extratension.

9. The greater the number of responses an individual gives in his free association to the Rorschach figures, the greater the amount of suggestibility he will display. Although Rorschach workers do not conceive of this score as such (except that few Rorschach responses might reflect constriction), Whipple (36) proposed that the number of associations to ink blots be indicative of the degree of suggestibility of an individual.

The aim of the present study was to test these hypotheses by determining the degree of relationship that exists between these specific Rorschach measures and the previously mentioned measures of suggestibility.

EXPERIMENTAL PROCEDURE

Subjects

Fifty students in the introductory psychology course at the State University of Iowa served as subjects. Only those *Ss* who gave a minimum of 15 scorable responses to all ten of the Rorschach cards, who were naive as far as the various aspects of the experiment were concerned, and who had no marked physical handicaps were used.[7] Thirty of the *Ss* were males and 20 were females. The age range was from 18 to 33 years with a mean of 21.8 years.

Apparatus

The apparatus used for the autokinetic sessions was an approximate duplicate of that described by Sherif (33) and used by more recent investigators (9, 10, 23, 34). A 6.3-v. radio dial lamp was enclosed in a black, lightproof box and was located 10 in. behind a .040-in. hole in a metal plate at the end of the box. A piece of opal flash glass was placed flush against the end of the box behind the metal plate to diffuse the light. A push-button switch was mounted on the arm of the *S*'s chair, which was approximately 18 feet from the light source. This button was used by *S* to signal the perception of movement. Five sec. following the pressing of the button by *S* the stimulus light was automatically turned off by an electronic time-delay relay. If the button was not pressed within 30 sec. the lamp was automatically shut off and "no movement" and a latency of 30 sec. were recorded. The light reappeared after an interval of 10 sec. (whether the button was pressed or not). Thus the maximum interval between trials was 40 sec. A Standard Electric Timer was used to record the latency between the onset of the light and *S*'s response. Both the timer and the recording sheets were kept in a large black box. The reading and the recording of the latencies and the estimates were done while *E* momentarily turned on a dim penlight.

A modification of the apparatus described by Hull (20, 21) was used to record body sway. It consisted of a 12-in. waxed paper polygraph and an electrically heated stylus that was connected by means of a thin fishline to the *S*'s head. This connection was accomplished without the *S*'s knowledge by having the fishline attached to the strap of a pair of goggles. These goggles were entirely blacked out, and when placed over the *S*'s eyes served both as a blindfold and as the means for recording sway. The recording of sway was made at full scale; no reduction pulleys being used as in Hull's system.

Procedure

In the first session the Rorschach was administered according to the procedure outlined by Beck (3) with the one exception that a card was removed at the end of five minutes if *S* had not returned it by that time.[8] This was done to allow for a manipulable number of responses for each *S* and to assure that the entire experiment would be completed within the three-hour limit that had been allotted for each *S*.

Immediately following the administration of the Rorschach, the "Ink-blot Suggestion Test" was given. Inasmuch as Eysenck and Furneaux (13) do not indicate what specific percepts were used in their study, it was decided to use two popular percepts from Beck's list (2) for the "common" responses and four responses from his *F*— tables as the "inapplicable" responses. In no case was a percept used in this test that had been given by *S* during the free association period to the Rorschach cards.[9] On showing each card *E* said, "Some people think this looks like a ———. Do you think it resembles that?"

Immediately following this, *E* and *S* proceeded to the room in which the autokinetic apparatus was located. Before entering the room, *S* was informed that the room would be completely dark but that to the right of the door there would be a

[7] Using these criteria, six *Ss*, of 56 volunteers, were excused from taking part in the experiment. Four gave fewer than 15 responses to the Rorschach; one had taken part in a preliminary study with the autokinetic apparatus; and one walked on crutches.

[8] It was necessary to do this with only seven of the 50 *Ss*. Four *Ss* had one card removed, while three had two cards withdrawn.

[9] Those percepts used for the "inapplicable" responses, and the frequency of use, were: (*a*) all of Fig. IV (inverted) as a "bee" (50); (*b*) all of Fig. V (inverted) named a "cow's head with horns" (48); (*c*) all of Fig. VI named a "brush" (47); (*d*) all of Fig. VII named a "frog" (50); (*e*) all of Fig. III (except area D 2) as a "mouth" (1); (*f*) all of Fig. III as a "spider" (1); (*g*) all of Fig. V as a "beetle" (1); and (*h*) all of Fig. VI as a "butterfly" (2).

chair in which he was to sit. After *S* was seated in the room and had become dark adapted the instructions were read. Except for a few minor modifications they were the same as those used by Sherif (33). Probably the most important change was that *S* was informed that the light would move anywhere from one to fifteen inches.[10] Fifty trials were given, the reported amount of movement and the response latencies (time taken to indicate the perception of movement) being recorded for each trial.

After this was completed, *S* was excused from the room on the pretext that the apparatus needed adjustment. While *S* was gone *E* calculated the mean of the judgments, determined the *planted* mean (the mean of the estimates that *E* was to provide in the second situation) and listed all of the scores that he was to give in the retest situation. On the basis of preliminary experimentation, it was decided that the planted mean would deviate 50 per cent from the mean estimate made by *S* in the first situation. For 26 *Ss* (those *Ss* whose initial mean estimates were under five in.) this planted mean was higher than the original mean; for 24 *Ss* (those whose mean estimates were greater than five in.) it was lower. Thus, the attempt was made to influence a lowering of estimates in those *Ss* who gave relatively large responses in the first situation and to raise the scores of those whose estimates were relatively small.

The *S* was then recalled to the experimental room and was told that the same procedure would be followed as previously, with the exception that *E* would announce, on each trial, the actual distance traversed by the light. On odd-numbered trials *E* made his statements first and on even-numbered trials *S* made his estimates first. The *E*'s estimates were in whole numbers and scattered narrowly around the planted mean. After 50 estimates were obtained from the *S*, he was again excused from the room. At this point the postural-sway apparatus was set-up.

Upon returning to the experimental room, *S* was asked to stand with his back to the apparatus while the goggles were fitted. He then was moved until the heated stylus was approximately 2 in. from the back edge of the paper.[11] He was told to stand still, erect, and with his feet together. At the end of the half-minute (baseline) period, *E* suggested that *S* was falling forward. Except for minor modifications, the procedure employed was the same as that used by Hull (21, p. 50). The suggestion was given repeatedly throughout a period of two minutes, during which continuous records were taken of *S*'s bodily position.

[10] This was done to yield manipulable scores. A preliminary investigation with this apparatus indicated that the estimates might vary from microscopic to astronomic dimensions.

[11] Under these conditions not much room was allowed for large backward sway. Of the 22 *Ss* who displayed negative suggestibility, however, only one consistently fell far enough backwards to reach the edge of the paper. Thus, only in the case of this *S* is it likely that the measure of his maximum sway was foreshortened.

RESULTS

In analyzing the results the following indices of suggestibility were employed:

a. The mean amount of movement reported during the first autokinetic session.

b. The degree of correspondence between the *S*'s revised mean estimate of movement (in the second autokinetic session) and the planted mean. In those instances where the planted mean was higher than *S*'s initial mean, this autokinetic ratio was calculated as follows:

$$\frac{M_r - M_o}{M_p - M_o}$$

(where M_o stands for the original mean; M_r stands for the revised mean; and M_p stands for the planted mean.) In those instances where M_p was lower than the initial estimate (M_o), M_r was subtracted from M_o (in the numerator) and M_p was subtracted from M_o (in the denominator) to obtain the ratio. Values ranging from zero (no change in estimates and, therefore, a minimum amount of suggestibility) to values greater than one (indicating a large shift in estimate and a high degree of suggestibility) were theoretically possible.

The mean time required to indicate the perception of movement in the first autokinetic session.

d. The number of percepts accepted in the "Inkblot Suggestion Test."

e. Scoring on the Postural-Sway Test was as follows: A baseline was drawn midway between the highest and lowest points in the initial half-minute period, i.e., before the suggestion instructions were read. The maximum sway in centimeters, either forward or backward, from this baseline was taken as the sway score for the individual.

The Rorschach protocols were scored according to the criteria of Beck (2). Whenever a response (or a similar percept) was not in the form-quality norms of Beck, the tables of Hertz (19) were consulted. If no such response could be found in either set of norms, the judgment of an expert Rorschach worker was obtained.[12]

Inspection of the frequency distributions of the variables indicated that a number of them (both Rorschach and the other test scores) were skewed.[13] Although the calculation of a product-moment correlation does not assume normality of the distributions, the interpretation of such statistics when normality does not exist is somewhat ambiguous(27). To allow for a more meaningful

[12] Special thanks are due to Dr. A. L. Benton for his assistance with the task of scoring a number of the protocols.

[13] This was to be expected in the case of at least some of the Rorschach scores (17), in the latencies (22), and the Postural-Sway scores (12, 21, 26).

interpretation of the scores, the data (with the exception of the latencies) were transformed into normalized standard scores. The logarithms of the latencies were used. All of the correlations (product-moment) were run using these transformed scores.[14]

The correlation coefficients between the various Rorschach measures and the "criterion" measures (with the exception of the autokinetic ratio) are presented in Table 1. Table 2 contains the correlations between the Rorschach measures and the autokinetic ratio. The first column contains the data for the group in which the attempt was to raise the estimate, while in the second column the data are those for the group in which the attempt was to lower the estimate.

Of the 72 correlations that resulted, only three reached the 5 per cent level of confidence or beyond. Four additional correlations reached a significance level between 5 per cent and 10 per cent. Thus, the great majority of the correlations are no greater than zero and not much assistance in prediction is afforded. This proportion of significant correlations (circa 10 per cent) would be expected by chance if there were no relationship between the Rorschach measures and the other indices.

[14] For a rationale and a discussion of the problems involved and the applicability of using transformed scores, see Mueller (27).

DISCUSSION

One can advance several possible reasons for the general lack of significance to the findings.

TABLE 1

CORRELATION COEFFICIENTS BETWEEN RORSCHACH SCORES AND THE "CRITERION" MEASURES (EXCLUSIVE OF THE AUTOKINETIC RATIO) † *

CATEGORY	INK-BLOT SCORE	LATENCY	MEAN ESTIMATE (AUTOKINETIC)	SWAY
R	.02	−.03	.11	.26
CF	.04	.02	.06	.09
%CF	.01	.22	.04	.03
M	−.19	.20	−.04	−.05
%M	−.19	.18	−.13	.11
F%	.04	.02	.20	−.09
F+%	−.12	.22	.12	−.01
F+% (on colored cards)	.01	.19	.18	−.13
A%	.03	.21	−.03	.11
P	.14	−.27	.10	.29
%P	.15	−.23	.08	.07
Average Color Score **	−.19	−.10	.00	−.02

† The signs of these correlations were arranged so that a positive correlation indicates conformity with the original hypothesis and a negative sign, the converse.

* The following are the correlations required at the indicated confidence levels (*df* of 48): .23=10 per cent, .28=5 per cent, .33=2 per cent, .36=1 per cent.

** The correlations in this row are based on the data of 49 *Ss*. One *S* gave no responses with color as a component, thereby making his average color score indeterminate.

TABLE 2

CORRELATION COEFFICIENTS BETWEEN RORSCHACH SCORES AND THE AUTOKINETIC RATIO † *

CATEGORY	RAISED (*N*=26)	LOWERED (*N*=24)
R	−.24	.25
CF	−.14	.41
%CF	−.14	.32
M	.07	.23
%M	.10	.13
F%	.00	−.17
F+%	.11	.07
F+% (on colored cards)	.35	.00
A%	−.01	−.55
P	.06	−.01
%P	.12	−.26
Average Color Score	.00**	.16

† The signs of these correlations were arranged in the same fashion as those in Table 1.

* The following are the correlations required at the indicated levels of confidence: *df*=24: .32=10 per cent, .39=5 per cent, .46=2 per cent, .50=1 per cent; *df*=22: .34=10 per cent, .40=5 per cent, .47=2 per cent, .52=1 per cent.

** This correlation is based on the data of 25 *Ss* since one *S* gave no response with color as a component.

It may be that these Rorschach indices represent still another type of suggestibility which is unrelated to the types assessed by the measures used here. It is essential, therefore, that the term in the Rorschach literature be more precisely defined than it has been to date and that suitable criteria be found so that an adequate test might be made.

Attention might also be directed towards the weaknesses inherent in the use of samples from a college population to test conclusions derived from experience with psychopathologic groups. Conceivably, the inference of suggestibility from Rorschach data may be applicable only to such deviant populations and, if this is the case, no adequate test of the hypotheses were made in this study. Further research might be oriented towards the investigation of Rorschach correlates of suggestibility in other populations. It should be noted, however, that the personality characteristic of suggestibility is not considered to be typical of only deviant personalities—it is presumed to exist in members of the general population. Only Beck, among the Rorschach workers who discuss suggestibility, mentions it when discussing deviant populations but he does not indicate that it is found exclusively in such groups. If the inference of suggestibility is applicable only in the presence of pathology, certainly this should be specified.

It is entirely possible, despite the statements of Rorschach, Beck, and Bell, that Rorschach test-measures do not correlate significantly with any test of suggestibility. The findings of the present study, within the limits imposed by its characteristics, indicate that this is a tenable hypothesis. Whether it applies to other Rorschach measures, other measures of suggestibility, and/or other samples, however, is of course an empirical question.

The relative homogeneity of the sample, as well as the restrictions in the ranges of some of the scores produced by the instructions (in the autokinetic situation), the setting of an arbitrary minimum number of Rorschach responses, and the withdrawing of some of the cards after five minutes probably lowered the obtained correlations somewhat. The use of a more heterogeneous sample, without the restrictions imposed in the present investigation, conceivably might lead to more significant findings.

In line with the statements of Rorschach workers of the need for considering the interrelationships between various components of the test, multiple correlations (or other similar technique) might profitably be employed. This, of course, would require specifying the variables to be used as well as employing a large sample (14, p. 455).

The findings of the present study would have definite implications for Rorschach practice in regard to the use of the term suggestibility. It is fairly obvious that whatever personality variable (or variables) is being assessed by the Rorschach measures utilized here, no satisfactory relationship exists with proposed measures of suggestibility. Whether they are related to other operational measures is a question which could be answered by other studies. For the time at least, the inference of suggestibility from these Rorschach indices is unwarranted.

SUMMARY

The purpose of the present study was to investigate the relationship between specific Rorschach test-measures presumably reflecting the personality variable of suggestibility and other, the majority of which had been previously used, measures of the same psychological characteristic.

Fifty undergraduate students in the introductory psychology course at the State University of Iowa were given the Rorschach test and then the following measures of suggestibility were obtained:

a. The mean amount of movement reported in 50 trials in an autokinetic situation.

b. The degree to which *S*'s estimates of movement reported in the autokinetic situation changed as a result of suggestions offered by *E* during a second situation.

c. The mean time required by *S* to indicate the perception of movement in the initial autokinetic session.

d. The number of percepts accepted by *S* in an "Ink-blot Suggestion Test," patterned after that of Eysenck and Furneaux (13).

e. The maximum amount of body-sway during the recitation of the suggestion that the subject was falling forward.

1. The findings indicated no consistent relationship between the Rorschach measures and the other indices of suggestibility.

2. Possible reasons for the lack of relationship between the Rorschach measures and the other measures were advanced and discussed.

3. It was concluded that the inference of suggestibility from the Rorschach indices used in this study is unwarranted at the present time.

REFERENCES

1. BECK, S. J. The Rorschach test as applied to a feeble-minded group. *Arch. Psychol., N. Y.*, 1932, **21**, No. 136.
2. BECK, S. J. *Rorschach's test.* Vol. I, *Basic processes.* New York: Grune & Stratton, 1944.
3. BECK, S. J. *Rorschach's test.* Vol. II, *A variety of personality pictures.* New York: Grune & Stratton, 1946.
4. BELL, J. E. *Projective techniques.* New York: Longmans, Green, 1948.
5. BENJAMIN, J. D., & EBAUGH, F. G. The diagnostic validity of the Rorschach test. *Amer. J. Psychiat.*, 1938, **94**, 1163–1178.
6. BIRD, C. Suggestions and suggestibility, a bibliography. *Psychol. Bull.*, 1939, **36**, 264–283.
7. BIRD, C. *Social psychology.* New York: Appleton-Century, 1940.
8. BLEULER, E. *Textbook of psychiatry* (trans. by Brill). New York: Macmillan, 1944.
9. BOVARD, E. W., JR. Social norms and the individual. *J. abnorm. soc. Psychol.*, 1948, **43**, 62–69.
10. BRAY, D. W. The prediction of behavior from two attitude scales. *J. abnorm. soc. Psychol.*, 1950, **45**, 64–84.
11. COFFIN, T. E. Some conditions of suggestion and suggestibility: a study of certain attitudinal and situational factors influencing the process of suggestion. *Psychol. Monogr.*, 1941, **53**, No. 4 (Whole No. 241).
12. EYSENCK, H. J. *Dimensions of personality.* London: Paul, Trench, Trubner, 1947.
13. EYSENCK, H. J., & FURNEAUX, W. D. Primary and secondary suggestibility: an experimental and statistical study. *J. exp. Psychol.*, 1945, **35**, 485–503.
14. GARRETT, H. E. *Statistics in psychology and education.* (2nd Ed.) New York: Longmans, Green, 1938.
15. GRIMES, F. V. An experimental analysis of the nature of suggestibility and of its relation to other psychological factors. *Stud. Psychol. Psychiat. Cathol. Univer. Amer.*, 1948, **7**, No. 4.
16. HERTZ, MARGUERITE R. The popular response factor in the Rorschach scoring. *J. Psychol.*, 1938, **8**, 100–121.
17. HERTZ, MARGUERITE R. Validity of the Rorschach method. *Amer. J. Orthopsychiat.*, 1941, **11**, 512–519.
18. HERTZ, MARGUERITE R. Rorschach twenty years after. *Psychol. Bull.*, 1942, **39**, 529–572.
19. HERTZ, MARGUERITE R. *Frequency tables to be used in scoring responses to the Rorschach ink-blot test.* (3rd Ed.) Cleveland: Western Reserve Univer., 1946.
20. HULL, C. L. Quantitative methods of investigating waking suggestion. *J. abnorm. soc. Psychol.*, 1929, **24**, 153–169.
21. HULL, C. L. *Hypnosis and suggestibility.* New York: Appleton-Century, 1933.
22. JENKINS, T. N. Facilitation and inhibition. *Arch. Psychol. N. Y.*, 1926, No. 86.
23. KELMAN, H. C. Effects of success and failure on "suggestibility" in the autokinetic situation. *J. abnorm. soc. Psychol.*, 1950, **45**, 267–285.
24. KLOPFER, B., & KELLEY, D. M. *The Rorschach technique.* Yonkers-on-Hudson, New York: World Book Co., 1942.
25. LEVY, D. M., & BECK, S. J. The Rorschach test in manic-depressive psychosis. *Amer. J. Orthopsychiat.*, 1934, **4**, 31–42.
26. MCCURDY, H. G. An experimental study of waking postural suggestion. *J. exp. Psychol.*, 1948, **38**, 250–256.
27. MUELLER, C. G. Numerical transformations in the analysis of experimental data. *Psychol. Bull.*, 1949, **46**, 198–221.
28. MURPHY, G., MURPHY, LOIS B., & NEWCOMB, T. M. *Experimental social psychology.* (Rev. Ed.) New York: Harper, 1937.
29. RORSCHACH, H. *Psychodiagnostics.* (Trans. by Lemkau and Kronenberg.) (2nd Ed.) New York: Grune and Stratton, 1942.
30. ROTTER, J. B. The present status of the Rorschach in clinical and experimental procedures. *J. Pers.*, 1948, **16**, 304–311.
31. SARGENT, HELEN D. Projective methods: Their origins, theory, and application in personality research. *Psychol. Bull.*, 1945, **42**, 257–293.
32. SCHNEIDER, L. I. Rorschach validation: Some methodological aspects. *Psychol. Bull.*, 1950, **47**, 493–508.
33. SHERIF, M. A study of some social factors in perception. *Arch. Psychol. N. Y.*, 1935, No. 187.
34. THURSTONE, L. L. *A factorial study of perception.* Chicago: Univer. of Chicago Press, 1944.
35. VARVEL, W. A. Suggestions toward experimental validation of the Rorschach test. *Bull. Menninger Clin.*, 1937, **1**, 220–226.
36. WHIPPLE, G. M. *Manual of mental and physical tests.* Baltimore: Warwick and York, 1910.
37. WILLIAMS, M. An experimental study of intellectual control under stress and associated Rorschach factors. *J. consult. Psychol.*, 1947, **11**, 21–29.

Received October 8, 1951.

CONDITIONING AND EXTINCTION OF THE GALVANIC SKIN RESPONSE AS A FUNCTION OF ANXIETY[1]

M. E. BITTERMAN AND WAYNE H. HOLTZMAN
The University of Texas

In recent years considerable interest has been shown in the relation between clinical indices of anxiety and behavior in conditioning situations. Welch and co-workers (16, 19, 20) have reported a number of experiments with the galvanic skin response in which children and adults diagnosed as pathologically anxious were found to condition more readily than normal controls. These experiments were undertaken in the hope of discovering an objective measure of anxiety and the papers in which the empirical relationships are reported do not deal at any length with the problem of theoretical interpretation. Comparable results were reported by Taylor (18), who studied the conditioning and extinction of the eyelid reflex in high and low anxiety groups selected from a student population on the basis of responses to 60 items from the Minnesota Multiphasic Personality Inventory (MMPI) which were judged to reflect anxiety. The high group conditioned more readily than the low and tended to extinguish less readily, although the latter difference was not statistically significant. According to Taylor, these results may be deduced from the theory of Hull (9) if anxiety is regarded as a secondary drive (14) since, in Hull's terms, ${}_{s}E_{R}$ (response evocation potential) is a multiplicative function of ${}_{s}H_{R}$ (habit strength) and D (drive). It is possible, as Taylor notes, that a difference in ${}_{s}H_{R}$ may be responsible,[2] but her data provide no basis for separating the two variables. A still more recent experiment by Hilgard, Jones, and Kaplan (7) failed to show a significant relation between Taylor's MMPI index and simple eyelid conditioning, but in a second phase of the study, which involved both reinforced and unreinforced stimuli, two indices of differentiation were found to be negatively correlated with the psychometric criterion. The results of the experiment are interpreted in rough cognitive terms; anxious subjects are said to react more in terms of their own apprehensions than in terms of carefully discriminated environmental objects and relationships.[3]

As Hilgard and co-workers have noted, their approach to the problem does not suggest a relation between anxiety and simple conditioning, but the theoretical position of Taylor does lead to the expectation of such a relationship. Although extreme groups (such as psychiatric samples vs. normals, or the upper and lower ends of the distribution of Taylor scores) have been clearly differentiated in terms of simple conditioning measures, the Hilgard experiment with a more homogeneous sample produced negative results. The failure of a significant relationship to appear in the latter experiment may, however, be a function of the insensitivity of the Taylor index. The present investigation, which was planned independently of the other recent studies, provides further evidence on this question. It was designed to permit a comparison of simple conditioning scores with the results of a more extensive clinical analysis of a homogeneous group of normal males. It was undertaken as a pilot study of the applicability of con-

[1] This research was conducted under terms of contract AF 33(038)-13887 between USAF and the University of Texas in coordination with the Department of Clinical Psychology, USAF School of Aviation Medicine, Randolph Field, Texas. We are indebted to R. Kaiser, J. Krauskopf, D. Ogdon, P. Reed, and G. Wilson, graduate students at the University of Texas, who assisted in the conduct of the experiment here reported.

[2] If the high anxiety group is considered to be more reactive than the low, the unconditioned stimulus will produce a stronger drive-state and its termination will provide greater need-reduction for the former group; in Hullian theory the development of habit strength is assumed to be a direct function of the amount of need-reduction.

[3] Similar findings, interpreted in Hullian terms, have been reported by Rosenbaum (15) for size discrimination when threat of strong shock was used to motivate rapid choice. High and low anxiety groups, selected from a student population on the basis of the Taylor index and from a psychiatric population on the basis of clinical diagnosis, performed differently, the high group showing a flatter generalization gradient (less precise discrimination) in each case.

ditioning procedures to the psychiatric screening of specialized Air Force personnel (17).

METHOD

The *S*s of the experiment were 40 nonveteran, unmarried male students, between 18 and 25 years of age, recruited from beginning classes in engineering at the University of Texas. They were informed that the study in which they were being asked to participate was sponsored by the Air Force and the request for their services was based primarily on an appeal to patriotic motives, but as an additional incentive, a token payment of five dollars was promised each *S* upon conclusion of the experiment. No information was given concerning the specific purpose of the research.

The experimental design called for four sessions—three individual and one group—which were scheduled at intervals of three to four weeks. The first session was devoted to the conditioning and extinction of the GSR, and the second to an individual Rorschach and brief interview. The third was a group session devoted to the administration of the MMPI, and, finally, each *S* was studied individually in a stress situation. The methods and procedures employed in each stage of the work will be described.

Conditioning and Extinction of the GSR

The conditioned stimulus (CS) employed in the experiment was a 5-sec., 1000-cycle tone of low intensity produced by a Hewlett-Packard audio-oscillator and led directly, without amplification, to a small speaker attached to *S*'s chair. The unconditioned stimulus (US) was a 3.5-ma. shock (from a 700-v. supply through a high series resistance) to the right wrist of *S*. The series resistance was so high that shock intensity remained fairly constant despite variations in skin resistance. Presentation of stimuli was automatically controlled. The duration of the shock was 1 sec., and its onset coincided with the termination of the CS; intertrial intervals ranged from 30 to 70 secs. The presentation of stimuli on each trial was recorded on one channel of a two-channel Brush ink-writing oscillograph (Model BL 202). A sample tracing may be seen in Fig. 1.

Skin resistance was measured between the palmar and dorsal surfaces of *S*'s left hand. Sponge-rubber electrodes saturated with a saline solution filled plastic cups (1½ in. in diameter) which were strapped to the surfaces of the hand. The rubber pads made pressure contact with zinc plates set into the back of the plastic cups and to these plates leads were soldered which ran to an external balancing and scaling device, a Brush D.C. amplifier (Model BL 932), and finally to the second channel of the recorder.[4]

The GSR session was conducted in a dimly illuminated soundproofed room. *S* was seated in a leather upholstered chair with his arms relaxed upon its broad armrests. The skin on the dorsal and palmar surfaces of the left hand and the wrist of the right hand was thoroughly cleaned and the GSR and shocking electrodes were attached. The apparatus, which was located in an adjoining room, from which *S* could be observed by *E*, was adjusted. *S* was informed that "various physiological measurements" would be made and that the experiment involved the administration of electric shock. He was asked to find a comfortable position in the chair and to sit as quietly as possible until the end of the session was announced. The *E* then entered the apparatus room and the experiment was begun. The first five trials constituted an adaptation series in which the CS was not reinforced. During the next ten trials—the conditioning series—each presentation of the CS was followed by the US. An extinction series, consisting of 20 unreinforced presentations of the CS, followed the conditioning trials.

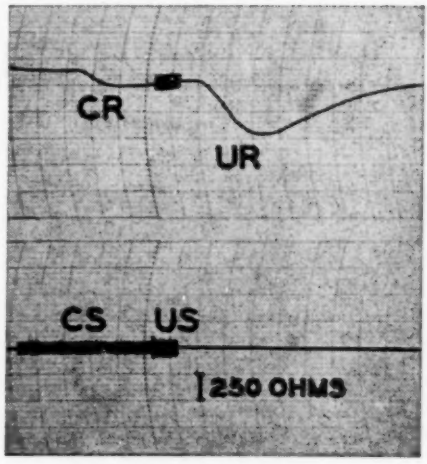

FIG. 1. SAMPLE RECORDING OF SKIN RESISTANCE

CS, conditioned stimulus; US, unconditioned stimulus; CR, conditioned response; UR, unconditioned response.

Interview and Rorschach Test

The second session was devoted principally to administration of the individual Rorschach, but in order to establish rapport and to obtain some background information a brief interview was conducted prior to the administration of the test. Conversation centered about *S*'s family relationships, occupational objectives, extracurricular interests, and social activities. At the conclusion of the interview the examiner rated *S*'s manifest anxiety on a five-point scale, and the rating was made again following the testing period.

The administration of the Rorschach was preceded by instructions similar to those recommended by Klopfer and Kelley (10). A standard inquiry was conducted in each case, and where certain

[4] A detailed description of the apparatus is presented elsewhere (1).

determinants were conspicuously absent, *S*'s limits of acceptance were tested in the usual manner.

MMPI

In the third phase of the experiment the MMPI was administered in accordance with the standard procedure outlined in the test manual (6). The *S*s were tested in two groups of approximately 20 each. Responses were recorded on IBM answer sheets and the blanks were machine scored.

The Stress Situation

The fourth and final phase of the experiment was designed to provide direct measures of performance under stressful conditions. The procedure employed was similar in many respects to that devised by Williams (22) in the evaluation of Rorschach factors related to intellectual control. No attempt was made to duplicate Williams' stress situation in every detail, although its major features were incorporated into the present plan. This phase of the work consisted of three parts: (*a*) the control period; (*b*) the stress period; and (*c*) the poststress interview. The three periods followed each other in rapid succession, the entire series requiring approximately one hour. Changes in test performance and other behavior from the control period to the stress period provided some indication of *S*'s ability to maintain control under tension and to recover control following the momentary disequilibrium induced by stress. The poststress interview was designed to permit further observations of the recovery process; *S* was asked to report on his reactions to the stress situation and an attempt was made to reduce any feelings of hostility which might have developed.

Three tests—a digit symbol substitution test, a size constancy test, and a manual steadiness test—were administered, in both the control period and the stress period. A brief description of each follows:

Digit Symbol. The digit symbol subtest of the Wechsler-Bellevue Intelligence Scale, Form I, as modified by Williams (22) was employed. By repetition of the last 43 items, the number of test items was increased from 67 to 110.

Size Constancy. The apparatus for this test consisted of two white equilateral triangles, each mounted on a black wooden background 12 x 20 in. These units were placed on a long table in such a way that one triangle was 6 ft. away from *S* and the other was 12 ft. away, behind and slightly to one side of the first. The nearer triangle, each side of which was 4.75 in. long, served as the standard. The distant triangle could be varied in size by *E* in stepwise fashion, *S* being instructed to say "Stop" when the two triangles appeared equal in size.

Steadiness. The apparatus employed was a modification of a standard device. The task of *S* was to hold a stylus fitted with a brass wire tip (.0625 in. in diameter) in a hole (.189 in. in diameter) cut out of a thin brass plate (4 x 8.5 in.) mounted at an angle of 45° from the vertical. While performing the task *S* stood before the plate which was 45 in. above the floor. The stylus was molded in such a way that it could not be completely withdrawn from the hole; when not in use it was lodged in a conical holder fitted to the underside of the plate. The apparatus was so wired that each contact of stylus and plate registered on a counter, and so that, when *E* desired, an electric shock could automatically be administered to *S* through a pair of electrodes attached to his left wrist.

Procedure. The experiment was conducted in two rooms, one for the control period and one for the stress period and poststress interview. The latter room, which is illustrated in Fig. 2, was divided into two sections by a partition into which were set two one-way viewing screens (half-silvered mirrors). A two-way public address system permitted verbal communication between the two sections of the room.

Control Period. The *S* entered the first room and was greeted cordially by *E* who introduced him to the observer (*O*). The *E* says, "Won't you please be seated here. Today we are going to take some more psychological tests, only these will be somewhat different from those that you have already taken." (*E* takes seat at table opposite *S*. *O* sits to one side and observes *S*.)

Digit Symbol. "You are about to be given a test which is an index of your general intelligence. The test which has been placed before you is known as the *Digit Symbol Test.* It is one frequently used in many intelligence test batteries. You will be given a number of trials on this test in which we are interested in seeing how high a score you can reach." (The standard instructions for the Wechsler-Bellevue Digit Symbol Test follow. Six trials of 90-sec. duration, separated by 30-sec. rest intervals, are given.)

Size Constancy. "Now for the next test. Would you please be seated over here in the little chair. Try not to move off of the chalk marks on the floor. You will notice in front of you, at different distances, two white triangles on black backgrounds." (*E* seats himself behind the variable triangle.) "Notice that I can vary the size of this one so that it becomes larger or smaller. Now I am going to set this one so that it appears obviously larger than the triangle which is nearer to you. Now I am slowly going to move this one so that it gets smaller and smaller, and when it appears to you that both triangles are equal in size, I want you to say 'Stop.'" (*E* moves triangle until S says "Stop.") "Now I am going to set this one so that it appears obviously smaller; again I am going to move it and again when it appears to you that both triangles are equal in size, just say 'Stop.'" (*E* moves triangle until *S* says "Stop.") "Now I am going to repeat this several times, some going up and some coming down. Each time they appear to you to be equal, just say 'Stop.' Are you ready?" (*E* makes a total of 8 determinations, half in the ascending and half in the descending order.)

Steadiness. "Now for the third test. Step over here and place your toes on the black line on the floor. What I want you to do is to see if you can hold the stylus in the center of the hole so that the metal tip does not make contact with the brass

plate; for every time it does, an error is recorded." (*E* demonstrates, showing *S* that errors will be recorded if the stylus is pulled out or pushed in too far as well as when it touches the brass plate.) "Remember, try to keep the metal tip from touching the sides of the brass hole. When I give the signal, withdraw the metal tip about one inch from its resting place and hold the stylus there until I say 'Stop.' Three seconds after the 'Go' signal, after you have had a chance to steady the tip in the hole, the error count will begin. When I say 'Stop,' the error count will stop and you will replace the stylus in the holder. Are you ready? Go." (Three 15-sec. trials are given with a 30-sec. rest interval between trials.)

overhead lights go out and two photoflood lamps in reflectors directed upon *S* are turned on. At the same time, the recorded instructions begin.)

"You are sitting in front of what appears to be a mirror but which in reality is a one-way screen. You are now being observed by a number of psychologists on the other side of the screen who are taking notes and continuous photographs of all your reactions throughout the remainder of this experiment. All directions are to be followed implicitly. Rest your arm attached to the electrodes on the table and keep it there from now on. You will notice that the electrodes on your arm are connected to the panel before you." (Pause . . . white panel light is snapped on.) "The white light which has

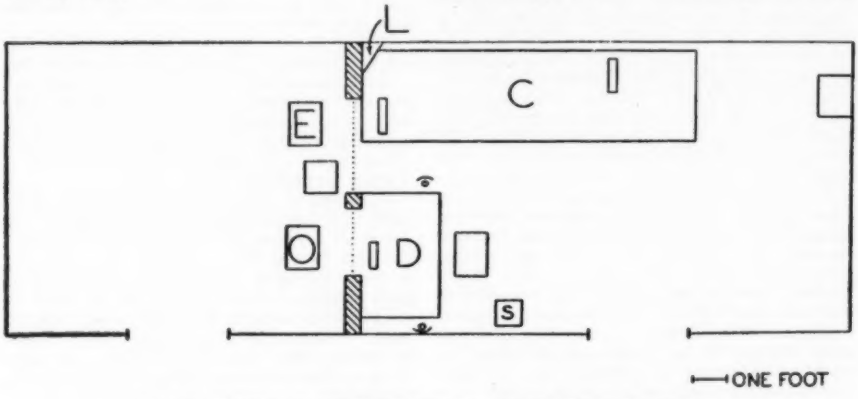

FIG. 2. DIAGRAM OF THE ROOM EMPLOYED FOR THE STRESS SITUATION.

The *S* worked in the section to the right. (*D*) designates the table at which he sat while occupied with the digit symbol test. The table is flanked by two photoflood lamps in reflectors. The small rectangle on the table shows the position of the signal lamps. (*S*) shows the position of the steadiness apparatus. (*C*) is a long table on which the constancy apparatus was placed, the two rectangles designating the positions of the standard (right) and variable (left). For this test the *S* sat in the chair indicated by the rectangle in the upper right-hand corner of the diagram. (*E*) and (*O*) show the chairs of the experimenter and observer. The position of the recording apparatus is shown by the square between their chairs. The dotted lines represent one-way mirrors. (*L*) shows the position of the speaker through which instructions reached the subject.

"Now if you will just be seated, there will be a three-minute rest period before the next experiment." (*O* takes his position behind the viewing screen in the room used for the stress period.)

Stress Period. "Now will you come with me please to another room where we shall proceed with the next experiment." (*E* leads *S* to the stress room.) "Please be seated in this chair facing this table. The rest of the experiment will be conducted in this room. Hold out your left (or right) hand. I am now placing a set of electrodes on your wrist, the function of which will soon be explained." (*E* attaches electrodes leading to shocking apparatus.) "I shall now leave the room. All further instructions will be given to you through a loudspeaker. Save any questions for the end of the experiment." (*E* leaves the room and joins *O* behind the viewing screen in the adjoining compartment. *S* is left alone for two full minutes. Suddenly the regular

just gone on indicates that our shock apparatus has been turned on. You are connected to that apparatus. During the following period you may receive a strong electric shock whenever the observers feel that your test performance is not up to our standards." (Red light is snapped on.) *"Whenever the red light goes on, you are not meeting our standards and you are in danger of being shocked . . .* like this!" (Standard shock, 3.5 ma., of 1-sec. duration is administered through the electrodes by *E*. Red light is turned off.)

Digit Symbol. "Now pick up a pencil and look at the test blank before you. It is the same test you took just a few minutes ago. Your instructions are the same as before. At the signal 'Go' work as fast as you can until you are told to stop. Concentrate on your work. Remember, your reactions are being observed and continually photographed, and *you will be shocked whenever your test perform-*

ance falls below our standards. Get set for trial one. Go!" (After 5 sec. the red light is snapped on. At 20 sec. a standard shock of 1-sec. duration is administered and the red light is turned off. At 30 sec. the red light is turned on again. The second shock of this trial is administered at 70 sec. Ten sec. after the standard shock the red light is again snapped on and it remains on for the rest of the trial. All trials are of 90-sec. duration.) "Stop! Place the test you have just completed on the right side of the table. You will be called upon shortly to repeat the test. Remember, you are being observed and continually photographed, and you will be shocked whenever your test performance is below our standards." (Silence is maintained for 40 sec.) "Get set for trial two. Go." (As on trial one, the red light is snapped on after 5 sec. At 30 sec. the standard shock is administered and the light is turned off. After 40 sec., the red light is turned on and remains on for the duration of the trial. After trial two the above instructions are repeated and the third and last trial is given. After 5 sec. the red light is snapped on. At 85 sec. the standard shock is administered. The light, turned off with the onset of shock, remains off for the duration of the trial.) "Now rise and go sit down in the small chair in the right rear of the room. Do not try to remove the electrodes from your hand; the wires are long enough to reach across the room."

Size Constancy. "You will notice that, as before, the rear triangle is larger than the one near you. In a few seconds the large triangle will slowly become smaller. When they appear equal in size, you are to say 'Stop.' The rear triangle will then be set so that it is smaller than the one near you. At the signal 'Ready' it will begin to get larger. Again, when the two triangles appear equal in size you will say 'Stop.' Several trials, both up and down, will be given. Each time the rear triangle will begin to move at the signal 'Ready.' And each time they appear to be equal you will say 'Stop.' Get set for the first trial." (Eight trials, four ascending and four descending, are given with a 30-sec. rest interval after each trial.) "Now rise and proceed toward the front of the room."

Steadiness. "Place your toes on the black line on the floor in front of the stylus device. Do not remove the electrodes from your hand; leave this hand hanging by your side. With your other hand, grasp the stylus which is in its holder." (Slight pause.) "At the signal 'Go' withdraw the metal tip from the holder about one inch. Three seconds after the 'Go' signal, the error count will begin. Hold the stylus there until the signal 'Stop.' Remember, try to keep the metal tip from touching the sides of the hole. Every time a contact is made, an error will be recorded. Get set. Go!" (Three trials of 15-sec. duration are given with a 20-sec. rest interval between trials.) "Stop! Replace the stylus. You may let your arm rest, but remain standing." (Pause for 20 sec.) "All right, when I give the signal, withdraw the metal tip as before and hold it there until I say 'Stop.' *This time when a contact is made, an error will be recorded and you will also receive a shock.* Get set. Go!" (Three trials of 15-sec. duration are given, with 20-sec. rest intervals as before.) "Stop! Replace the stylus. Remain standing where you are. At the next signal 'Go,' remove the metal tip from the holder again. *This time try to see how long you can hold the stylus without touching the sides of the hole.* Every time a contact is made, an error will be recorded and you will receive a shock. The error count will begin three seconds after the 'Go' signal. *You may terminate this trial any time you wish after the 'Go' signal by simply saying 'Stop.' Having said 'Stop,' you cannot start again.*" (Slight pause.) "In a previous group of high school students, a maximum time of two minutes and 17 seconds was made by one of the subjects. See if you can break this record. *The subject in this group who holds the stylus in the hole the longest time and with the least amount of errors will receive a prize of five dollars.*" (Slight pause.) "To terminate the trial, simply say 'Stop.' Are you ready? Go!" (One min. after the "Go" signal the shock intensity is increased from 2.5 to 3.5 ma. If *S* persists at the end of 2 min., the shock intensity is increased to 6.5 ma. After 3 min. the trial is terminated by *E*.) "Stop. You have been timed at three minutes. If anyone else in this group reaches three minutes the one with the lowest error score will receive the five dollar prize." (Instructions stop.)

Poststress Interview. The poststress interview was conducted in an informal manner. While *E* was questioning *S*, *O* recorded their behavior from behind the one-way screen. Although the interview was relatively unstructured, nine specific questions were asked by *E* in the course of conversation with *S*:

1. How well do you think you did?
2. What especially was there about it that made you feel this way?
3. Did you feel like you were "on the spot"?
4. Did you mind the shock much?
5. Did you think you were going to get more shock?
6. Did you find the experiment interesting?
7. Do you feel that you might do better at some other time?
8. Do you think you did better in this room (stress) or back in the other room?
9. Do you have any other comments?

Before *S* left the room he was asked not to tell any of the other *S*s about his experience in the room. Such instruction was given after all phases of the experiment and the importance of conformity was repeatedly stressed.

The Behavior Rating Schedule. A behavior rating schedule was used by *O* to evaluate the behavior of each *S* in the stress situation. Although the schedule generally resembled the one employed by Freeman and co-workers (4) in a study of the stress interview, it was designed to fit the specific needs of the present experiment. Ratings of all *S*s were made by the same observer. The behavior of the *S* was evaluated in relation to each of the following items: (*a*) degree of tension arousal, (*b*) speed of tension arousal, (*c*) physical poise, (*d*) degree of attention or problem orientation, (*e*) shock response, (*f*) speed of tension reduction, (*g*) response

to red warning light, (*h*) memory for instructions, (*i*) recovery as reflected by conversation in the post-stress interview. A five-point scale was used in rating each item. The first four items were rated also in the control period.

RESULTS

The Criterion of Anxiety

The second, third, and fourth phases of the experiment were designed to provide a criterion of anxiety with which to compare conditioning and extinction behavior, and the first step in the analysis of results was to derive such a criterion. The problem was approached by a team of four clinically trained psychologists (H, Ka, Og, and W) in the following manner:

Ka and W, working independently, evaluated the objective scores obtained in the stress situation. Their attention was centered primarily on changes in digit symbol score from control to stress period, general level of steadiness, changes in steadiness from control to stress period, the variability of constancy scores obtained in the stress period, and the duration of the persistence trial with the steadiness apparatus. On the basis of these considerations, the anxiety manifested by each of the 37 *Ss* for whom complete data were available[5] was rated on a five-point scale (step 5 indicating high anxiety and step 1 low anxiety). The correlation between the two sets of ratings, corrected for errors of grouping (5), was .68.

Og and Ka, working independently with the behavior rating schedule, made over-all ratings of the anxiety of each *S* on a five-point scale. The correlation between the resulting sets, corrected for errors of grouping, was .91.

H and Og independently examined the Rorschach data and made ratings of susceptibility to anxiety on the five-point scale. Of the many aspects of Rorschach performance which have been regarded as indicators of anxiety (2), the following were emphasized in the ratings: form quality (22), Elizur's index (3), Holtzman's impulsiveness index (8), the number of responses,[6] and card rejections. In this case, the agreement between the two sets of ratings was not high, the coefficient of correlation (corrected for errors of grouping) being .39.

In the next stage of the analysis, W and Og, working together, scrutinized the six sets of ratings already described, referring frequently to the original data upon which they were based, and agreed upon a new over-all rating for each individual (again on a five-point scale). H and Ka conferred in the same way to produce a comparable set of ratings. The correlation between the two sets, corrected for grouping errors, was .55.

The final ratings which served as the clinical criterion of anxiety were arrived at in a conference of all four raters. The earlier ratings were reviewed and discrepancies were resolved on the basis of re-examination of the raw data. In this conference MMPI profiles were considered for the first time, special attention being given to the Taylor index (18) and the neurotic triad (13). The distribution of ratings is shown in Table 1.

TABLE 1

THE DISTRIBUTION OF CRITERION RATINGS

SCALE STEP	*N*
1. (low anxiety)	1
2.	5
3.	13
4.	13
5. (high anxiety)	5
Total	37

On the basis of these ratings the *Ss* were divided into two criterion groups approximately equal in size. The 19 *Ss* with ratings between 1 and 3 were designated the low anxiety group and the 18 *Ss* with ratings of 4 or 5 comprised the high anxiety group. The relation between anxiety and conditioning was studied in terms of a comparison of the performance of these two groups in the conditioning situation.

Performance of High and Low Anxiety Groups in the Conditioning Situation

The behavior of the *Ss* in the conditioning situation is expressed in terms of changes in skin resistance. The first question to be decided in dealing with the results was the

[5] Of the original 40 *Ss*, one became seriously ill and two left school during the course of the experiment.

[6] The response frequencies obtained from the *Ss* of this experiment were sharply limited although every effort was made to encourage productivity. Over half the records contained less than 15 responses and only four contained more than 30. This result seems less surprising when it is recalled that the *Ss* were students of engineering.

appropriateness of change in resistance as a unit of measurement. Lacey and Siegel (11, 12) have suggested that the unit of measurement employed in galvanic studies should meet two important criteria: (*a*) it should be normally distributed and (*b*) it should be uncorrelated with the basal level of resistance. These investigators report that resistance change satisfies neither of these criteria, but that conductance provides a more satisfactory unit of measurement. Since the range of initial resistance levels was attenuated in the present experiment, it was not considered wise to accept these conclusions without further check. For this reason, product-moment correlations were computed between basal resistance level and deflections expressed both in resistance and in conductance. These correlations were computed for responses to four critical stimuli: the first tone, first shock, tenth shock, and the tone of the first extinction trial. From Table 2 it may be seen that the resistance measure was not significantly correlated with level, while the conductance measure did correlate signifi-

TABLE 2

CORRELATIONS BETWEEN RESPONSES TO VARIOUS STIMULI (EXPRESSED IN RESISTANCE AND CONDUCTANCE) AND BASAL LEVELS

STIMULUS	UNIT	
	RESISTANCE	CONDUCTANCE
First Tone	.25	−.18
First Shock	.23	−.13
Tenth Shock	−.04	−.27
Tone (First Extinction Trial)	.06	−.52*

* $p < .01$.

cantly at one point. For this reason the resistance measure was used in presenting the results of the experiment. Since the distributions of both units were somewhat skewed, all statistical tests were made by nonparametric methods.

In Fig. 3, median response to the conditioned stimulus ("tone" response) for paired trials during all three phases of the procedure—adaptation, conditioning, and extinction—is plotted separately for the high and low anxiety groups derived from the clinical ratings discussed in the preceding section.[7]

[7] It may be well to emphasize at this point that the conditioning data and the criterion data were taken by different investigators. The individual GSR recordings were analyzed independently in order to avoid the possibility of contamination, and the scores were later grouped in accordance with the criterion ratings.

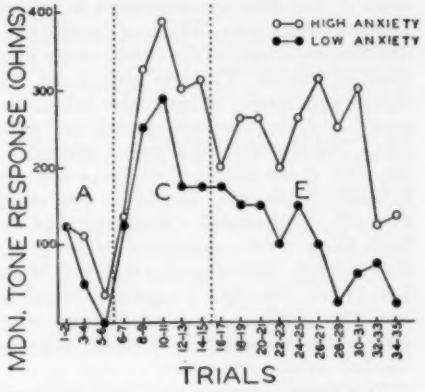

FIG. 3. ADAPTATION, CONDITIONING, AND EXTINCTION OF THE "TONE" RESPONSE IN HIGH AND LOW ANXIETY GROUPS

The three phases of the procedure are indicated by (A), (C), and (E).

The curves for both groups seem to show the same general trends. The initial (unconditioned) response to tone disappears to a great extent during the unreinforced adaptation trials; during the conditioning trials the tone response grows in negatively accelerated fashion, and then it tends to decline once more during the extinction series. The significance of these trends was tested by comparing the distributions of scores at various points on the curves by Wilcoxon's nonparametric test for paired deviates (21). The adaptation effect was tested in terms of the difference between responses on trials 1–2 and trials 5–6, and the difference for both groups was significant at the 1 per cent level of confidence. Both groups also conditioned significantly ($p < .01$) as indicated by the differences between points 6–7 and 14–15. During the extinction series, however, different trends appeared in the two groups. The low group showed significant extinction as measured by the difference between responses on trials 16–17 and trials 34–35 ($p < .01$), but the decline for the high anxiety group was not significant ($p > .05$).

Inspection of Fig. 3 shows that the curve for the high anxiety group is generally higher than that for the low group. The signifi-

cance of this difference was tested at various stages of training by Wilcoxon's method for unpaired replicates.[8] The results of the tests were as follows: The two groups did not differ significantly during the adaptation period or during the early conditioning trials (6–9), but the level of response attained by the high group during trials 10–15 was significantly higher than that of the low group ($p<.05$). The over-all responsiveness of the high group during extinction (trials 16–35) was also significantly greater than that of the low group ($p<.01$). In general, therefore, the high anxiety group conditioned more readily and extinguished less readily than did the low anxiety group.

The pattern of response to the conditioned stimulus was often found to be quite complex during the extinction trials. The initial deflection produced by the CS—which has been referred to as "tone" response—frequently was followed within a few seconds by a second deflection which coincided temporally with the unconditioned response to shock during the conditioning trials; in some records as many as three or four deflections appeared. In order to take these supplementary responses into account, a second index was computed which is referred to as "total" response. It was defined as the difference between the resistance level at the onset of the CS and the lowest point reached by the pen in the next eight seconds.

In Fig. 4 the course of extinction is plotted in terms of total response for high and low anxiety groups. The results for total response during extinction are the same as those described for tone response. The curve for the low group falls rapidly while that for the high group maintains its initial level. Wilcoxon's test for paired deviates showed no significant difference in the responsiveness of the high group on trials 1–2 as compared with 19–20, while for the low anxiety group this difference was significant beyond the 1 per cent level of confidence.

In order to determine whether the obtained differences in conditioning and extinction were a function of the differential effective-

[8] Since this method requires that the number of cases in the two groups be equal, the data for one *S* in in the low anxiety group were eliminated from consideration. This *S* was chosen at random from among the 13 with anxiety ratings of 3.

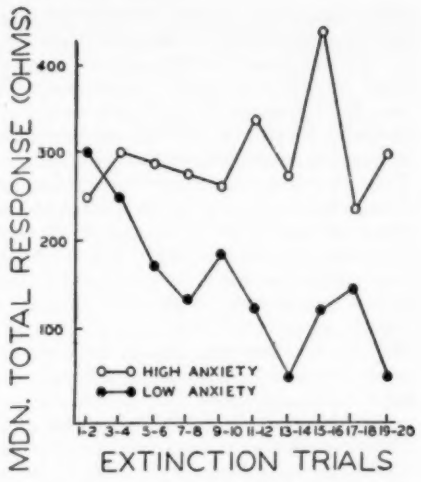

FIG. 4. EXTINCTION OF THE "TOTAL" RESPONSE IN HIGH AND LOW ANXIETY GROUPS

ness of the unconditioned stimulus for the two groups, the magnitude of the UR to shock was measured on each of the ten conditioning trials. Median values for the two groups are plotted in Fig. 5. Although the basal levels of the groups did not differ significantly at any point in the experiment, the

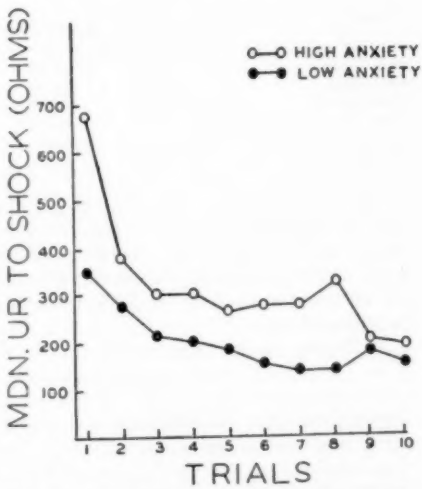

FIG. 5. ADAPTATION OF THE UNCONDITIONED RESPONSE TO SHOCK IN HIGH AND LOW ANXIETY GROUPS

high group showed a significantly greater unconditioned response on the first trial (Wilcoxon's test for unpaired replicates, $p<.05$). These *Ss* adapted rapidly, however, and the difference between the groups failed to be significant after the first trial. It seems unlikely, therefore, that all of the results of the experiment can be accounted for in terms of the differential effectiveness of the shock.

Upon completion of the analysis described above, the *Ss* were divided into two new groups on the basis of Taylor's MMPI index. Eighteen *Ss* with Taylor scores ranging from 12 to 40 constituted the new high anxiety group and 19 *Ss* with Taylor scores ranging from 2 to 11 comprised the new low anxiety group. While a comparison of the new groups in terms of conditioning and extinction measures showed mean differences in the same direction as those obtained from a comparison of the old groups, these differences were considerably short of statistical significance. Neither were the Taylor scores significantly related to the clinical ratings upon which the old grouping was based ($\chi^2=2.21$, $p>.05$). These results are in accord with findings of Hilgard, Jones, and Kaplan (7). Taylor's (18) positive findings may be attributed to the fact that she worked with extreme ends of the distribution of scores on the MMPI items.

SUMMARY AND CONCLUSIONS

The relation between a clinical index of anxiety and behavior in a conditioning situation was studied in a homogeneous group of normal males. Thirty-seven university students were rated for susceptibility to anxiety on the basis of psychometric indices and performance in a laboratory stress situation, and in terms of these ratings they were divided into two groups. The galvanic skin response to shock conditioned more readily and extinguished less readily in the high anxiety group than in the low. In view of the homogeneity of the sample, this experiment represents a most stringent test of the hypothesis that anxiety is related to the rate of conditioning and extinction. The results suggest that conditioning scores may be sensitive enough to be of practical value in the psychiatric screening of specialized military personnel.

REFERENCES

1. BITTERMAN, M. E., & HOLTZMAN, W. H. *Development of psychiatric screening of flying personnel. III. Conditioning and extinction of the galvanic skin response in relation to clinical evidence of anxiety.* USAF School of Aviation Medicine, Randolph Field, Texas, 1952.
2. BENTON, A. L. The experimental validation of the Rorschach test. *Brit. J. Med. Psychol.*, 1950, **23**, 44–58.
3. ELIZUR, A. Content analysis of the Rorschach with regard to anxiety and hostility. *Ror. Res. Exch. and J. proj. Tech.*, 1949, **13**, 247–284.
4. FREEMAN, G. L., MANSON, G. E., KATZOFF, E. T., & PATHMAN, J. H. The stress interview. *J. abnorm. soc. Psychol.*, 1942, **37**, 427–447.
5. GUILFORD, J. P. *Fundamental statistics in psychology and education.* (2nd Ed.) New York: McGraw-Hill, 1950.
6. HATHAWAY, S. R., & MCKINLEY, J. C. *Manual of the Minnesota Multiphasic Personality Inventory.* (Rev. Ed.) New York: Psychological Corporation, 1943.
7. HILGARD, E. R., JONES, L. V., & KAPLAN, S. J. Conditioned discrimination as related to anxiety. *J. exp. Psychol.*, in press.
8. HOLTZMAN, W. H. Validation studies of the Rorschach test: impulsiveness in the normal superior adult. *J. clin. Psychol.*, 1950, **6**, 348–351.
9. HULL, C. L. *Principles of behavior.* New York: D. Appleton-Century, 1943.
10. KLOPFER, B., & KELLEY, D. M. *The Rorschach test.* New York: World Book Co., 1946.
11. LACEY, O. L. An analysis of the appropriate unit for use in the measurement of galvanic skin resistance. *J. exp. Psychol.*, 1947, **37**, 449–457.
12. LACEY, O. L., & SIEGEL, P. S. An analysis of the unit of measurement of the galvanic skin response. *J. exp. Psychol.*, 1949, **39**, 122–127.
13. MEEHL, P. E. Profile analysis of the M.M.P.I. in differential diagnosis. *J. appl. Psychol.*, 1946, **30**, 517–524.
14. MOWRER, O. H. A stimulus response analysis of anxiety and its role as a reinforcing agent. *Psychol. Rev.*, 1939, **46**, 553–565.
15. ROSENBAUM, B. Stimulus generalization as a function of clinical anxiety. *Amer. Psychologist*, 1950, **5**, 327.
16. SCHIFF, E., DOUGAN, C., & WELCH, L. The conditioned PGR and the EEG as indicators of anxiety. *J. abnorm. soc. Psychol.*, 1949, **44**, 549–552.
17. SELLS, S. B. *A research program on the psychiatric selection of flying personnel.* Report No. 1. *Methodological introduction and experimental design.* USAF School of Aviation Medicine, Randolph Field, Texas, 1951.
18. TAYLOR, JANET A. The relationship of anxiety to the conditioned eyelid response. *J. exp. Psychol.*, 1951, **41**, 81–92.
19. WELCH, L., & KUBIS, J. Conditioned PGR (psychogalvanic response) in states of pathological anxiety. *J. nerv. ment. Dis.*, 1947, **105**, 372–381.
20. WELCH, L., & KUBIS, J. The effect of anxiety on the conditioning rate and stability of the PGR. *J. Psychol.*, 1947, **23**, 83–91.
21. WILCOXON, F. *Some rapid approximate statistical procedures.* New York: American Cyanamid Co., 1949.
22. WILLIAMS, M. An experimental study of intellectual control under stress and associated Rorschach factors. *J. consult. Psychol.*, 1947, **11**, 21–29.

Received October 12, 1951.

ACCEPTANCE BY OTHERS AND ITS RELATION TO ACCEPTANCE OF SELF AND OTHERS [1]

CHARLES J. McINTYRE
The Pennsylvania State College

ONE of the postulates of Rogers' (2) theory of personality and behavior states, in effect, that an individual's acceptance of himself is positively and significantly correlated with his acceptance of others. This view is supported experimentally by the study of Sheerer (3) in which statements relating to attitudes toward self and others made by clients undergoing psychotherapy were rated by judges, and the Pearson product-moment *r*'s between the attitudes were found to be in the +.50's and +.60's.

Recently, Phillips (1) has converted the self-others attitudes as reported by Sheerer into simple statements to form a questionnaire of 50 items, with 25 items referring to self attitudes and 25 items referring to attitudes toward others. Typical of the items referring to attitudes toward self are:

My own decisions regarding problems I face do not turn out to be good ones.

I find that I feel the need to excuse or apologize for my behavior.

Items referring to attitude toward others are typified by:

When others make an error in my presence I am almost certain to point it out to them.

I feel it hard to take a general interest in the activities of some of my friends.

The items are presented randomly and answered by the respondents on a five-point scale—A, B, C, D, E—where A means "rarely or almost never true for me," B means "sometimes but infrequently true for me," C means "occasionally true for me," D means "very often true for me," and E means "true for me all or almost all of the time." The items are arbitrarily weighted with one point for A, two for B, etc., to five points for E. Thus, scores for acceptance of self and acceptance of others are obtained. Correlations between the two scales from +.51 to +.74 were obtained when the questionnaire was administered anonymously to high school and college students. Phillips found a test-retest reliability after a five-day lapse of +.84 for the questions regarding self and +.82 for the questions regarding others. When reliability was studied, the questionnaires were not administered anonymously, and the self-other correlation is not reported.

While Phillips' results support Rogers' postulate of a positive relationship between attitudes toward self and others, they do not throw any light upon Rogers' statement that " . . . the person who accepts himself will, because of this self-acceptance, have better interpersonal relations with others" (2, p. 520). It was to investigate this relationship that the present study was undertaken. That is, are individuals who better accept themselves and others better accepted by others?

PROCEDURE

Subjects

Subjects were male college students living in a large dormitory. All 315 of the men living on the first, second, and third floors of the dormitory responded to a sociometric questionnaire, described below, but the results were counted only for the 224 men who were white, second-semester freshmen who had been living on the same floor since the beginning of the first semester. The other men were eliminated inasmuch as they might receive a low sociometric rating because of short residence, race, age difference, or other factors clearly not pertinent to this study.

Method

All *Ss* were given a sociometric questionnaire to which they responded by designating, in order of preference, up to eight men living on their floor with whom:

1. They would most like to spend an evening of recreation.

2. They would most like to room if they had a choice.

By assigning a value of from eight for first choice to one for eighth choice, it was possible to compute a total "acceptance" score for each man. For example, if a man were chosen first by two men (8×2), third by three men (6×3), and eighth by one man (1×1), his total score would be 35.

The *Ss* were then ranked by floors from most to least accepted, i.e., highest to lowest score. Then the upper and lower 25 per cent of each group were given the Phillips attitude-toward-self-and-others questionnaire.

An identical sociometric questionnaire had been given to those *Ss* living on one floor of the dormitory two months after the beginning of the first semester and approximately six months prior to this study. It was possible, therefore, to compute a reliability coefficient for the sociometric questionnaire.

[1] The author wishes to express his appreciation to William G. Shipman for his assistance in formulating this study and in collecting the data, and to Dr. William U. Snyder for his helpful comments.

RESULTS

1. In the opinion of the resident dormitory advisers, who were graduate students in psychology and who had been associated closely with the *Ss* for eight months, the sociometric questionnaire very successfully discriminated those who were highly accepted by the group from those who were poorly accepted.

2. The test-retest coefficient of reliability for the sociometric questionnaire was +.65 as computed by the Pearson product-moment method.

3. The Pearson product-moment correlation between the acceptance-of-self and acceptance-of-others scale of the Phillips questionnaire was +.46.

4. The differences between the mean scores of the highly and poorly accepted groups on both the acceptance-of-self and acceptance-of-others scales were such as to occur by chance at least 50 times out of 100. Such differences as occurred between the mean scores of the two groups clearly could be attributed to chance.

5. The mean of the highly and poorly accepted groups combined was found for each scale. Then in each group the number of cases lying above the mean was compared by chi square with the number which could be expected to fall above the mean by chance, i.e., 50 per cent.

On the acceptance-of-self scale, the probability was 70 in 100 that the differences between the observed and expected frequency occurred by chance.

On the acceptance-of-others scale, the probability was 90 in 100 that the differences between the observed and expected frequency occurred by chance.

DISCUSSION

The coefficient of reliability of +.65 is difficult to evaluate since it may reflect variance due to a change in group structure as well as the less than perfect ability of the questionnaire to yield consistent results. Thus, while reliability is ordinarily a measure of the consistency of the measurement, in this case some inconsistency is inherent in the function being measured. The writer is well acquainted with the social grouping from which this reliability measure was taken, and there seems to be little doubt that much of the variance is due to an actual change in the group structure.

The correlation of +.46 between the acceptance-of-self and acceptance-of-others scale of the Phillips questionnaire, while somewhat lower than those reported by Phillips, reflects a significant relationship between acceptance of self and others as measured by this questionnaire.

The results of this study in no way support Rogers' view that better interpersonal relations are a function of better attitudes toward self and others. At any rate, one's attitudes toward self and others as measured by this questionnaire were not found to be related to acceptance by others, which would seem to be an integral part of interpersonal relations.

However, the negative results of this experiment may be largely the result of the method and Phillips' questionnaire.

With respect to method, the questionnaire could not be administered anonymously, and this may have had the effect of discouraging honest answers. In addition, the questionnaire itself is rather superficial; the *Ss* easily could give "good" answers rather than correct ones, and there is no way that the dishonesty could be detected.

CONCLUSION

1. The sociometric device used appears to be a fairly reliable method with a good face validity in discriminating individuals in a dormitory situation who are highly or poorly accepted in the group in which they live.

2. The results on Phillips' questionnaire tend to support Rogers' postulate that attitudes toward self and others are positively and significantly correlated.

3. The results are ambiguous with respect to Rogers' hypothesis that better attitudes toward self and others are positively related to better interpersonal relations inasmuch as the results might be attributed at least as easily to the method of the experiment as to the incorrectness of the hypothesis.

REFERENCES

1. PHILLIPS, E. L. Attitudes toward self and others: a brief questionnaire report. *J. consult. Psychol.*, 1951, 15, 79–81.
2. ROGERS, C. R. *Client-centered therapy*. Boston: Houghton Mifflin, 1951.
3. SHEERER, ELIZABETH T. An analysis of the relationship between acceptance of and respect for self and acceptance of and respect for others in ten counseling cases. *J. consult. Psychol.*, 1949, 13, 169–175.

Received October 20, 1951.

RATINGS ON TRAITS MEASURED BY A FACTORED PERSONALITY INVENTORY

JOHN B. CARROLL
Harvard University

THIS is a report on correlations obtained between self-ratings, peer-ratings, and test scores on five traits which the Guilford Inventory of Factors STDCR (3) purports to measure. The chief concern in the present study was to see whether the factorially defined traits of Guilford's inventory could be reliably measured and separated by rating-scale techniques. The results can be interpreted, if one so desires, as giving evidence on the validity of Guilford's inventory, but Ellis (2) has pointed out certain limitations in the use of interpersonal ratings as a method of validating personality inventories. Guilford and Martin (6) have already reported a minor study of the validity of the Inventory of Factors STDCR; their results may be compared with the data reported here. In any case, the validation of Guilford's inventory is of secondary interest in the present context.

METHOD

According to Guilford's test manual, "the factors S, T, D, C, and R taken together probably cover the area of personality generally encompassed by the concept of introversion-extraversion." The test consists of 175 questions of the Yes-?-No type, and yields scores on five traits identified in factor analysis studies (4, 5). These traits are described in the test manual as follows:

S—social introversion, as exhibited in shyness and tendencies to withdraw from social contacts.
T—thinking introversion, an inclination to meditative thinking, philosophizing, and analyzing one's self and others.
D—depression, including feelings of unworthiness and guilt.
C—cycloid tendencies, as shown in strong emotional reactions, fluctuations in mood, and tendency toward flightiness or instability.
R—rhathymia, a happy-go-lucky or carefree disposition; liveliness and impulsiveness.

Some of the planning of this study and the collection of data were carried out as a laboratory exercise in the tests and measurements course of the Army Specialized Training Program curriculum in personnel psychology at the University of Chicago in the fall of 1943. Approximately 125 Army enlisted men, with a median age of about 25, were enrolled in this course. Most of the men had had some previous college training. At the time the data were collected, these men had been quartered together (usually six to a room) in the university dormitories for four months. It was felt that conditions were particularly favorable for securing reliable interpersonal ratings, since the men were living in close association, were all taking an identical course of study, and saw each other constantly every day.

The first problem was to construct rating scales of the traits set forth in the manual of Guilford's Inventory of Factors STDCR. The men were given the descriptions of these traits and told to develop graphic rating scales for each. This assignment was given as if it were merely an exercise in the construction of rating scales; the men were not told until later that the trait descriptions had been extracted from a personality test, and that they would eventually take this test and rate each other on the scales. Through cooperative effort and with the help of the writer, who was the instructor in the course, the rating scales reproduced in Table 1 were finally developed. Some difficulty was experienced in deciding on the nature of the traits at the opposite ends of the scale, since in each case the trait description in the test manual gave a clue on the nature of only one end of the scale. The instructor tried to decide what the opposite ends of the scales were like by examination of the scoring of the items. It was, of course, crucially important that the rating scales be based on the same traits as the personality test, but there was no way of knowing whether this ideal was achieved.

After the rating scales had been constructed, each man did three things, in order: (*a*) he rated himself on each of the five rating scales; (*b*) he rated each of his five roommates; and (*c*) he filled out the Guilford Inventory of Factors STDCR. These were accomplished at a single session; the Guilford test was given last so that the ratings would not be biased by it. A few men had fewer than five roommates; in these cases the peer-ratings were obtained in such a way that each man received ratings from five other men who were best acquainted with him. It should be added here that the data were handled with as much anonymity as possible; no man could identify the scores of others, nor could he tell with any certainty which of his roommates gave him any particular rating.

After the session in which the data were obtained, the purpose of the study was divulged and the men were assigned the task of collating, analyzing, and interpreting the results. The writer, however, has completely reanalyzed the data, all based on the 110 men for whom complete data were available.

TABLE 1

RATING SCALE OF FIVE PERSONALITY TRAITS

(Each graphic rating scale contained eleven points; points 2, 4, 6, 8, 10 were designated as indicated below.)

(S) I. Does the individual seek or avoid social contacts?
- 2. Goes out of way to seek social contacts.
- 4. Frequently seeks social contacts; seems to enjoy being with people.
- 6. Neither seeks nor avoids social contact.
- 8. Frequently avoids social contacts; seems to prefer being alone.
- 10. Goes out of his way to avoid social contacts.

(T) II. Is the individual inclined to meditative thinking, introspection and analyzing the thoughts and actions of himself and others?
- 2. Rarely engages in meditation or introspection about himself or others.
- 4. Occasionally engages in meditative activities and introspection.
- 6. Engages in a moderate amount of meditative activity.
- 8. Indulges in considerable meditation about himself and others.
- 10. Is much given to meditative activities, continually analyzing the thoughts and actions of himself and others.

(D) III. Does the person tend to be depressed about himself and his achievements and to have feelings of unworthiness and guilt?
- 2. Is cheerfully confident in spite of troubles; has marked self esteem and belief in his own worth and the value of his achievements.
- 4. Usually satisfied and confident, but with occasional feelings of depression and dissatisfaction with his achievements.
- 6. Alternates between moderate optimism and moderate pessimism about his achievements, and often is in doubt about them.
- 8. Frequently has the "blues"; is self-accusing. Does not often feel satisfied with his actions.
- 10. Is continually depressed, and has strong sense of unworthiness and guilt.

(C) IV. How well does the individual control his emotions when frustrated?
- 2. Is very stable, has very well-balanced emotional control, shows emotional reactions only when very seriously frustrated; rarely moody.
- 4. Usually well-balanced but shows emotion when moderately frustrated; seldom moody.
- 6. Fairly good emotional control but has occasional strong emotions; occasionally moody for no apparent reason.
- 8. Finds it difficult to control emotions when frustrated; may be frequently moody, fluctuating rather readily from excitement to depression.
- 10. Shows strong emotional reactions; has marked fluctuations in mood, and is markedly disturbed by the slightest frustration. Often shows emotion without any apparent reason.

(R) V. Does the individual tend to take a carefree or or a serious attitude toward things in general?
- 2. Very serious; concerned about the future and the results of his actions; usually slow and deliberate; cautious; seldom takes an attitude of indifference toward anything brought to his attention; serious about things which don't really matter.
- 4. Takes a serious attitude toward things, but occasionally shows indifference to incidental matters.
- 6. Serious about things which matter, but indifferent and mildly carefree about incidental matters.
- 8. Can be serious about some things of great importance to him, but usually happy-go-lucky and not worried about outcomes.
- 10. Is happy-go-lucky and carefree; is hardly ever truly serious; doesn't take anything seriously; is impulsive; not concerned about the future or the results of his actions; irresponsible.

RESULTS

Central Tendencies, Dispersions, and Reliabilities of Ratings and Test Scores

Table 2 shows the means and standard deviations of these measures, as well as the percentile rank values ($P.R._{\bar{x}}$) of the test means, based on graphical interpolation of the norms given in the test manual. The sample shows minor divergences from the norm group: with respect to traits D, C, and R it is on the whole slightly less "depressed," "cycloid," and "happy-go-lucky" than the norm group of 388 University of Nebraska students. Examination of the mean ratings, in conjunction with the designations on the rating scales, will show how the typical person is characterized (either by himself or by his roommates). There is one interesting difference between mean self-ratings and peer-ratings: the men tended to rate themselves as more meditative than their room-

mates thought them to be. On the whole, however, the differences do not appear to be of any practical significance. The ratings show considerable dispersion over the scale; the fact that the dispersions of mean peer-ratings are smaller than those of self-ratings is due to regression effects which come about through averaging of ratings of five roommates.

Data reported in the test manual show that the scores on the Inventory of Factors STDCR are highly reliable. In one sample, reliabilities of .92, .89, .91, .91, and .89 for factors S, T, D, C, and R, respectively, are reported. No further estimates of reliability of these test scores have been made in the present study.

The reliabilities of the peer-ratings on each of the five traits are shown in Table 2. Ranging from .664 to .806, they are regarded as sufficiently high for the purposes of the present study. The highest reliability was obtained for the trait called "social introversion": one might expect that among the five traits this is the one which is most overtly manifested and most easily observable by others. The lowest reliability was obtained for "thinking introversion," which clearly has to do with a type of subjective behavior not too likely to be evident to others. These reliabilities are in general distinctly higher than those reported by Guilford and Martin (6) for a somewhat similar series of ratings. They report reliabilities of .69, .01, .73, .38, and .54, respectively, for the five traits, based on 22 individuals, each rated by five acquaintances.

TABLE 2

DATA ON SELF-RATINGS, PEER-RATINGS, AND TEST SCORES ON TRAITS S, T, D, C, R

(N=110)

TRAIT	SELF-RATING		AVER. PEER-RATING			INVENTORY SCORE		
	MEAN	*SD*	MEAN	*SD*	r_{∞}*	MEAN	*SD*	*P.R.*$\bar{x}$**
S	4.94	1.43	4.76	1.30	.806	17.8	10.4	53
T	6.64	1.88	5.46	1.28	.664	35.7	10.2	49
D	4.70	1.50	4.11	1.20	.729	18.8	11.1	43
C	4.60	1.73	4.24	1.38	.746	24.1	11.5	37
R	5.40	1.72	5.23	1.50	.763	39.1	14.3	43

* r_{∞}=estimated reliability of the average of five ratings.

** *P.R.*$\bar{x}$=percentile rank value of mean, by graphical interpolation in norms given in the test manual.

The reliability of the self-ratings cannot be estimated from the present data, since only one rating per trait was made.

The reliability of the peer-ratings can be estimated by the use of analysis of variance techniques. We estimate the reliability of a *single* rating (i.e., by *one* rater) by the formula

$$r_{11}=1-\frac{\text{Within Individual Variance}}{\text{Total Variance}},$$

where the "within individual variance" is the estimate of the variance of the five roommates' ratings, and the "total variance" is the variance of all $5N$ ratings involved for a given trait. The reliability of the composite rating by five roommates (r_{55}) is then estimated from r_{11} by the Spearman-Brown prophecy formula. The soundness of this procedure for determining reliability of multiple ratings has been discussed by Ebel (1).

Intercorrelations Between Traits

To economize on space, all the intercorrelations between the fifteen variables analyzed here (self-ratings, mean peer-ratings, and inventory scores on each of five traits) are assembled in Table 3. The reader may wish to rearrange the data in order to see more clearly the form of the results which will now be discussed.

This study was designed to see whether the traits measured by the Guilford inventofy are independent, and whether their relative independence would be changed if they were measured by rating-scale techniques. Three sets of correlations in Table 3 may be considered: those between the five traits (*a*) as

TABLE 3

CORRELATIONS BETWEEN SELF-RATINGS (s), PEER-RATINGS (p), AND TEST SCORES (t) ON FIVE PERSONALITY TRAITS *
(N=110)

TRAIT		S			T			D			C			R		
		s	p	t	s	p	t	s	p	t	s	p	t	s	p	t
S	s	..	42	57	26	20	11	31	17	19	11	01	—01	—10	—28	—53
	p	.	..	37	02	37	08	08	49	10	08	19	—01	—31	—55	—38
	t	.	.	..	28	23	27	44	31	62	31	08	36	—15	—21	—69
T	s	.	.	.	..	39	65	32	09	42	21	06	26	—12	—17	—37
	p	.	*51*	.	.	..	32	40	38	15	21	16	04	—31	—56	—40
	t	.	.	*30*	.	.	..	25	11	57	20	15	47	—30	—20	—32
D	s	.	.	.	.	.	.	..	29	53	47	14	45	—04	—08	—29
	p	.	*64*	.	.	*55*	.	.	..	27	27	40	24	—30	—34	—25
	t	.	.	*68*	.	.	*63*	.	.	..	54	20	90	—12	03	—28
C	s	.	.	.	.	.	.	.	.	.	..	30	52	—06	—01	—13
	p	.	*25*	.	.	*23*	.	.	*54*	.	.	..	26	—07	07	05
	t	.	.	*39*	.	.	*52*	.	.	99	.	.	..	—04	16	06
R	s	.	.	.	.	.	.	.	.	.	.	.	.	..	56	34
	p	.	—*70*	.	.	—*78*	.	.	—*46*	.	.	*10*	.	.	..	45
	t	.	.	—*76*	.	.	—*36*	.	.	—31	.	.	07	.	.	..

* Decimal points omitted. Correlations appearing below the diagonal are the results of corrections for attenuation.

measured by the Guilford inventory, (*b*) as measured by self-ratings, and (*c*) as measured by peer-ratings. The mean absolute values of the correlations in each set are, respectively, .45, .20, and .35. Thus it would appear that the self-rated traits generally show most independence, while the traits represented by inventory scores show most dependence. It can be argued, however, that these results may be accounted for partly on the basis of differential reliability. Correction for attenuation would seem to be in order. In the absence of any estimate of the reliabilities of self-ratings, it was impossible to correct these correlations for attenuation, but the other sets could be corrected in this manner, as shown in Table 3.[1] The mean absolute values of the corrected correlations were .50 for the inventory scores and .48 for the peer-ratings. Thus, the peer-ratings now seem to be quite similar to the test scores with respect to the dependence or independence of the five traits. Whether the self-ratings would show as much over-all correlation as peer-ratings or test scores if corrections for attenuation could be applied, we cannot say. The fact that self-ratings were capable of having fairly high correlations with other variables suggests that their reliabilities are reasonably high, however—perhaps as high as those of the mean peer-ratings. We would therefore anticipate that correlations among self-ratings, when corrected for attenuation, would not be as high on the average as the corrected correlations between peer-ratings or between inventory scores. At any rate, neither the self-ratings, peer-ratings, nor the test scores can be said to measure five clearly independent traits, but the traits are somewhat more sharply separated in the ratings.

Relations Between Self-Ratings, Peer-Ratings, and Inventory Scores on a Given Trait

If we examine, in Table 3, the correlations between the three measurements for each trait, we find that the highest correlations (with r's ranging from .34 to .65) are generally those between self-ratings and personality test scores. Such a result might be anticipated on the ground that both self-ratings and personality test scores reflect an individual's own estimates of himself.

Peer-ratings and self-ratings are correlated, on corresponding traits, with zero-order correlations ranging from .29 to .56.

With some qualifications, to be discussed later, the correlations between the peer-ratings and the personality test scores may be considered as evidence for the "validity" of the latter. The zero-order correlations in this case range from .26 (for Trait C, cycloid tendencies) to .45 (for Trait R, rhathymia or "happy-go-luckiness"), and indicate only very moderate validity for the personality inventory.

Of major interest is the question of whether the traits defined by the personality inventory are separated in the same way as they are when measured by self-ratings and peer-ratings. This question could perhaps be answered most authoritatively by a factor analysis of the complete table of intercorrelations; it would then be possible to observe whether the three types of measurements cluster in roughly the same positions, according to the trait being measured, in a space of five factors. This is the result we should expect if there are truly five dimensions of personality covered by our data, and if the three types of measurement are factorially equivalent. As yet, however, the data have not been subjected to factor analysis.

A reasonably satisfactory procedure is to notice whether a given measurement correlates most highly with another measusement when the same trait is involved. For example, if we take self-rating trait S, we find that it correlates more highly with peer-rating trait S than with the other peer-rating traits. On the other hand, self-rating trait D correlates more highly with peer-rating trait T than it does with peer-rating trait D. If we proceed with this type of analysis, we gain an over-all impression that the highest correlations are found between corresponding measures of the same trait, but that there are a considerable number of exceptions.

Another type of analysis is the use of multiple regression. Each set of variables (self-ratings, peer-ratings, and test scores) can be used as a battery of predictors for each variable in the other two sets. If the five

[1] The reliabilities used for these corrections were, for the test scores, those reported by Guilford in his test manual, and for the peer-ratings, those presented in Table 2.

traits are structured in a comparable way regardless of how they are measured, then it should be found that the highest beta-weight for the prediction of any given variable is for the corresponding trait in the predictor variables. This analysis has been performed by the writer, but cannot be reported here in detail.[2] It will suffice to report the number of times each trait is predicted best by its corresponding trait in the predictor variables; there are six multiple correlations involved in each case. For Trait S, highest beta-weights are found on corresponding traits 5 times; for Trait T, 3 times; for Trait D, twice; for Trait C, 4 times; and for Trait R, 4 times. From these results it can be concluded that the five traits are fairly well separated from each other in a manner which prevails regardless of whether they are measured by a personality test, by self-rating, or by peer-ratings. On the other hand, there are certain cases of indistinctness and confusion. The confusions are most often found between Traits S (social introversion) and R (rhathymia) and between Traits D (depression) and C (cycloid tendencies).

DISCUSSION

The interpretation of the results reported here is limited, to some extent, by the fact that one cannot assert with confidence that the rating scales were constructed so as to measure what they were intended to measure, namely, the traits represented by the five scores on Guilford's Inventory of Factors STDCR. The only evidence on this point comes from the multiple regression of inventory scores on self-ratings. The results for Traits S (social introversion) and T (thinking introversion) appear to show that the graphic rating scales were properly constructed, since the self-ratings on the respective traits are the only variables requiring relatively high beta-weights in the prediction of the corresponding inventory scores. In the case of Traits D and C, the picture is not too clear, but it must be noted that these traits are not at all sharply separated by the inventory—their intercorrelation after correction for attenuation being .99. Possibly the rating scale for Trait C involves too much the notion of unstable emotions *in response to frustration*, but the correlation obtained between self-rating and inventory score on this trait, .52, is appreciably higher than the value of .34 reported by Guilford and Martin (6) with reference to a scale whose polarity is stated as follows: "Has frequent or radical changes in mood. . . . Very uniform in mood." Trait R is the one that shows the lowest zero-order correlation, .34, between self-rating and inventory score, but the writer cannot perceive any logical discrepancy between the statements used in the rating scale and the description of the trait presumably measured by the test. It is a reasonable interpretation that the self-rating on this scale is a valid measure of a trait very imperfectly measured by the test. This is inferred from two facts: (*a*) the self-ratings on Traits R and S are largely independent ($r = -.10$), but the corresponding inventory scores are rather highly correlated ($r = -.53$); and (*b*) in the prediction of inventory score on Trait R, the self-rating on Trait S (rather than R) contributes the greatest independent variance. Our best course is to conclude that the graphic rating scales were constructed so as to measure the traits described as those measured by the test, but that the latter measured those traits imperfectly and with unsatisfactory differentiation.

There is reason to think that the subjects accomplished the ratings and tests carefully, seriously, and honestly. They had little reason to falsify their responses, since the scores were to have no effect on their Army careers and since data were handled with relative anonymity. The fact that the responses were probably made honestly is an advantage in view of the purpose of this study, but it limits the extent to which one can generalize from these results to a situation where falsification would be more likely. Particularly, it qualifies any conclusions one might draw as to the "validity" of the personality questionnaire used in this study.

If one grants that self-ratings have satisfactory reliability, they seem to have more intrin-

[2] For the detailed statistical results order Document 3604 from the American Documentation Institute, 1719 N St., N.W., Washington 6, D. C., remitting $1.00 for microfilm (images 1 in. high on standard 35 mm. motion picture film) or $1.00 for photocopies (6 x 8 inches) readable without optical aid. These may also be obtained from the author on request, if the request is made within not too long a time after publication.

sic validity than peer-ratings. They give a broader picture of the individual's personality, because they are more independent. Peers see only the external, superficial aspects of an individual's personality, and they tend to rate in terms of certain stereotypes. For example, the correlation between roommates' ratings on Trait S (social introversion) and R (happy-go-luckiness) was —.55, while the corresponding correlation in the self-ratings was only —.10, from which we can infer that while roommates think that social extraversion is a sign of a carefree attitude, these traits are largely independent in the subjective view of a man rating himself. The scoring of the personality test, also, seems to have made the traits appear more correlated than they actually are. Likewise, the scoring of the personality test makes it appear that Trait T (thinking introversion) is more closely associated with depressive and cycloid tendencies than is the case if we accept the subjective impressions of people when they rate themselves. Even from the standpoint of peer-ratings, thinking introversion is relatively independent of depressive and cycloid traits.

The total significance of these results, then, seems to be that self-ratings give the clearest separation of the five traits presumably measured by the Guilford Inventory of Factors STDCR, that roommates' ratings give a fairly clear separation, and that the scoring keys of the inventory give the least clear separation.

SUMMARY

1. Self-ratings, peer-ratings, and personality test scores on five traits purported to be measured by the Guilford Inventory of Factors STDCR were obtained for 110 enlisted men enrolled in an Army Specialized Training Program in personnel psychology. The ratings were made on graphic rating scales specially constructed to correspond to the traits of the personality test.

2. The reliabilities of average ratings made by five roommates were satisfactory, ranging from .664 to .806, and appeared to depend to some extent on the overtness of the behavior being observed.

3. The five traits appeared to be more independent when self-ratings were analyzed; the trait interdependence was about the same when measured by peer-ratings and the personality test.

4. For any given trait, the correlation between self-rating and the personality test score was usually highest, probably because the same person made the appraisal; next highest, in general, was the correlation between self-rating and peer-rating. The correlations between corresponding traits on the personality test and in the peer-ratings were generally lowest, ranging from .26 to .46. With some reservations, it is inferred from these latter results that the "validity" of the Guilford Inventory of Factors STDCR is only very moderate. The highest correlations are not always found between corresponding traits, however, as is shown by a multiple-correlation analysis.

5. It is suggested that with respect to the five traits under study, self-ratings give a broader picture of the individual's personality, since roommates tend to rate in terms of superficial kinds of behavior and in terms of stereotypes, and the scoring keys of the personality inventory yield scores which are in many cases too highly dependent on each other.

REFERENCES

1. EBEL, R. L. Estimation of the reliability of ratings. *Psychometrika*, 1951, 16, 407–424.
2. ELLIS, A. The validity of personality questionnaires. *Psychol. Bull.*, 1946, 43, 385–440.
3. GUILFORD, J. P. *Inventory of factors STDCR.* Beverly Hills, Calif.: Sheridan Supply Co., 1940.
4. GUILFORD, J. P., & GUILFORD, RUTH B. Personality factors S, E, and M, and their measurement. *J. Psychol.*, 1936, 2, 109–127.
5. GUILFORD, J. P., & GUILFORD, RUTH B. Personality factors D, R, T, and A. *J. abnorm. soc. Psychol.*, 1939, 34, 21–36.
6. GUILFORD, J. P., & MARTIN, H. Age differences and sex differences in some introvertive and emotional traits. *J. gen. Psychol.*, 1944, 31, 219–229.

Received October 27, 1951.

STIMULUS GENERALIZATION AFTER EXTINCTION AND PUNISHMENT: AN EXPERIMENTAL STUDY OF DISPLACEMENT

F. R. BRUSH, R. R. BUSH, W. O. JENKINS, W. F. JOHN, AND J. W. M. WHITING

Harvard University [1]

STIMULUS generalization has been studied by a number of investigators (9), although only a single reference, that of Skinner's (10), has been made to stimulus generalization with responses such as bar-pressing by rats or pecking by pigeons. On the other hand, there is extensive literature on experimental extinction of such responses (6). No experimental studies known to the authors have been made of generalization of an instrumental response after extinction. A similar situation exists with respect to generalization after punishment; punishment of an instrumental response has been studied (3), but generalization after punishment has not, except for the work of Miller and Kraeling (8).

In spite of this lack of experimental data, there is considerable theoretical interest in the effects of extinction and punishment on subsequent generalization. In a provocative paper (7), Miller discussed the relation of generalization to psychoanalytic displacement and formulated a theory of generalization under conflict conditions. One of Miller's major conclusions was that if a response is inhibited by punishment so that conflict is produced, its strength will *increase* along a dissimilarity dimension to a maximum and then decrease to zero. The assumptions from which this deduction follows include generalization decrement with stimulus change for both the original and conflicting responses (approach and avoidance) and a steeper falling off of the generalization function for the conflicting (avoidance) behavior. A different prediction arises from the simple combination of the principles of punishment and generalization. Both treatments are followed by a decrement in response strength. Thus one might reasonably expect the curve of generalization under conflict conditions to be a monotonically decreasing function of stimulus dissimilarity. This prediction also follows from a mathematical model presented by Bush and Mosteller (2).

The experiments described in this paper were designed to answer some of the questions raised in the preceding paragraphs. A Skinner-type situation was used in which the strength of the conflicting response was not directly measured, but rather inferred from the decrement in the original (approach) behavior. Conflict was assumed to be present when response strength declined under extinction or punishment conditions. Generalization of an instrumental response was studied first after conditioning alone, next after experimental extinction following conditioning, and finally after the conditioned response had been suppressed by punishment.

METHOD

Subjects and Preliminary Training

Subjects (*Ss*) were five groups of adult pigeons trained to peck an illuminated frosted window on a 12-minute, aperiodic, partial reinforcement (APR) schedule (5). Four groups had undergone three previous training periods, each followed by an extinction. The fifth group was naive prior to this study. Before the generalization tests, all birds were run for one 108-minute training session a day for about a week. The birds were selected to represent a wide range of rates of responding. Drive was held constant by maintaining the birds at 80 per cent of original body weight throughout the experiment.

Apparatus

The major features of the apparatus used have been described previously (5). Briefly, four boxes were equipped with illuminated one-inch diameter windows to be pecked by the birds. A timer was connected to the food magazine so that the first peck after an average interval of 12 minutes was followed by presentation of food. The time intervals used were 8, 14, 12, 10, 16 minutes. Four birds were run simultaneously. Ink records of cumu-

[1] This research was a joint product of the Laboratories of Human Development and of Social Relations, Harvard University. One of us (R. R. B.) is greatly indebted to the Social Science Research Council and the National Research Council for granting him a Joint Fellowship in the Natural and Social Sciences that made his contribution to this work possible. The authors wish to thank the following individuals for advice: W. W. Lambert, N. E. Miller, K. C. Montgomery, F. Mosteller, R. R. Sears, and B. F. Skinner.

lative responses versus time and counter readings at 12-minute intervals were obtained.

The window to be pecked was illuminated by a small projector outside of the experimental box. The stimulus situation was varied by changing the size of the spot of light on the window.[2] This was accomplished by rotating a brass disc in which various sized holes were punched. The sizes of the several light spots in cm. diameter were as follows:

Cm. diam.	2.6	2.2	1.8	1.4	1.0	0.6	0.2
Code no.	+3	+2	+1	0	−1	−2	−3

The identifying code number is given under each diameter. It may be noted that there were three larger spots and three smaller spots with respect to the training spot (0).

Control Group

One group of four birds was trained to a stable rate of responding on the 12-minute APR schedule and then tested with each of the generalized stimuli for 12 minutes. No responses were reinforced during test periods. (This method was used by Skinner [10].) After each test, reconditioning with the training spot, 0, to the original level of responding took place. The order of presentation of the six test spots was different for each bird. All birds were given three presentations of each test stimulus.

Extinction Groups

Two groups of four pigeons were trained in the same way as the control group and then extinguished to a criterion of 10 minutes of no response. After a bird in Group I reached this criterion, a 12-minute test with one of the generalized stimuli was given. The training spot (0) was then re-presented until the criterion of extinction was again reached and then another test spot was presented for 12 minutes. This sequence was repeated until each of the six test spots had been presented twice.

Extinction Group II was included for two reasons: (*a*) it was suspected that the training with spot size −3 in Group I, prior to this experiment, biased the generalization results, and (*b*) tests with the standard stimulus (spot size 0) were not given under the same conditions as tests with the other stimuli so that the data for spot size 0 were not strictly comparable to the other test data. (This criticism of the procedure used for the first extinction group was made by F. Mosteller.) Hence Group II consisted of four experimentally naive birds which received tests with spot size 0 as follows: After the criterion of 10 minutes of no responding was met with spot 0, the projector disc was rotated to another spot 0 for 12 minutes. This procedure paralleled that followed for the six other tests.

An attempt was made to extinguish a third group of four birds to approximately 50 per cent of their maximum rates and then perform the generalization tests. However, it was found that it was very difficult to stop extinction at the 50 per cent rate and impossible to hold birds at this rate long enough to give the tests.

Punishment Groups

Three birds were given the same preliminary training as the Control and Extinction Groups. Shock was then introduced in the following fashion. Two heavy-duty B-batteries were wired to the grid floor of the box so that 67.5 volts could be administered through the birds' feet by a knife switch. A preparation of graphite mixed with machine oil was smeared on the birds' feet for conducting purposes. The shock was administered approximately according to the same 12-minute APR schedule as food. Each shock lasted for about three seconds and was presented immediately after a response or series of responses, but never in conjunction with food. A period of at least one minute of no shock always preceded and followed presentation of food.

A combination of food-shock training continued until responding stabilized at 50 per cent of the preshock level. Since there are local variations in rate of responding, a 50 per cent rate is difficult to maintain; therefore limits of 45 per cent to 55 per cent for 12-minute periods were employed. When such a rate was maintained, a test with one of the generalized stimuli was introduced for the following 12-minute period. During tests, neither shock nor food was presented. After a test the original shock-food conditions were reinstated until responding stabilized for another 12 minutes, after which another test was given. This procedure was followed until all seven spot sizes had been tested twice, in a different order for each of the three birds.

Two additional birds were trained as above, but with the frequency of shock increased so that response strength was reduced to an average of 3 per cent of its original level. The actual limits of responding ranged from 1 to 8 per cent prior to testing. One series of tests with the seven stimuli was given during which neither food nor shock was presented. The procedure paralleled that for the group maintained at 50 per cent of original response strength. Following these tests, frequency of shock was increased for the two birds until a criterion of 10 minutes of nonresponse was reached. Food training was continued. Each time the criterion was met it was followed by a 12-minute test with one of the seven stimuli. One such series of tests was given. The procedure paralleled that of Extinction Group II.

From casual observation, qualitative reactions to shock early in training included (*a*) a tense jerk followed by an apparent increase in breathing rate, (*b*) temporary thrashing or freezing, and (*c*) brief cessation of pecking followed by recovery. Later in shock training when the rate of pecking stabilized, all these reactions seemed to be less pronounced. Rate was somewhat more variable with shock than without.

[2] This method was suggested to the writers by G. Heise and K. Safford. It was also used by Skinner, Heise, and Safford in a demonstration experiment.

RESULTS

Control Group

In the Control Group data, Table 1, the numbers of responses in 12 minutes shown are means of the three tests on each bird and each spot, except for spot 0. The scores for spot 0 are means of 18 12-minute periods for each bird; these data were obtained without rotating the projector disc and so are not strictly comparable to the scores for the six other spot sizes (see discussion of procedure for the second extinction group). Nevertheless, those scores were taken as reasonable estimates of the pretest rates of responding. Also in Table 1 are shown the mean ratios of the test scores to the number of responses in the immediately preceding 12 minutes with spot 0; average ratios for the four birds are shown for each spot. In Fig. 1 the mean ratio data are plotted.

It is apparent from Table 1 that all birds in the Control Group produced declining response strength as stimulus dissimilarity was increased. Of the 24 possible comparisons (six for each bird), there were only four reversals from the expected direction, and these were small in magnitude.

From Table 1 and Fig. 1, it is clear that the generalization curve is asymmetrical with respect to the training spot when diameter is used as the abscissa. While there is no a priori reason for expecting symmetry in such a plot, the training of the birds previous to this experiment may have contributed to the observed asymmetry. That previous training involved prolonged conditioning with a spot size closely approximating +3. When one plots the ratios of Table 1 against the logarithm of spot diameter, however, the curve is much more symmetrical.

Of interest is the comparison of the three tests of each bird in the Control Group with each spot size. An analysis of variance of the ratio data yielded *F* values that failed to reach the 5 per cent level for both birds and tests. This was true for each of the six generalized stimuli. Hence it was justified to consider the three tests on each bird as 12 independent observations of the response strength at each stimulus. This conclusion, with respect to bird-to-bird differences, was expected since ratios were used throughout. The conclusion with respect to test-to-test differences, however, indicates that negligible discrimination learning took place during the experiment.

An analysis was made of the data arranged according to order of presentation of the test stimuli regardless of size for each group. No significant progressive effects were found although there was a tendency for responding to the first test stimulus to be somewhat higher than that to the other stimuli. (This finding also held for the other groups discussed below.)

TABLE 1

CONTROL-GROUP RESPONSES IN 12 MINUTES DURING TESTS WITH GENERALIZED STIMULI

BIRD	SPOT SIZE *						
	−3	−2	−1	0	+1	+2	+3
9	85	213	551	718	762	380	363
	(.12)	(.32)	(.75)		(.90)	(.63)	(.58)
10	29	144	510	560	400	347	283
	(.10)	(.24)	(1.13)		(.86)	(.55)	(.55)
11	36	148	296	326	328	317	217
	(.08)	(.42)	(.83)		(1.02)	(1.00)	(.59)
12	15	93	143	200	192	132	120
	(.09)	(.41)	(.68)		(1.01)	(.69)	(.55)
Mean	41	150	375	451	420	294	246
	(.10)	(.35)	(.85)		(.95)	(.72)	(.56)

* The numbers in parentheses are ratios of test responses to the number of responses in the immediately preceding 12 min. with the training spot (0).

Extinction Groups

In the data for the Extinction Groups, Table 2, the means of raw responses in two

tests are shown for each bird and each spot size. Also shown are the ratios of these mean scores to the mean pre-extinction rate for each bird; average ratios for the eight birds are given. It will be recalled that Extinction Group II differed from Extinction Group I in that the former had had no training prior to this experiment (and thus no bias on spot size +3) and that tests were made with spot size 0 in exactly the same fashion as with the other spots.

The results for the two groups are essentially the same. The over-all level of responding is low, not exceeding 4 per cent of the pre-extinction rates. In both groups there is an apparent upswing in the curve at the larger spot sizes. However, an analysis of variance for each group and for both the raw and ratio scores failed to yield *F* values significant at the 5 per cent level of confidence for between-spots variance. This result does not mean that the data are inconsistent with an hypothesized small increase along the similarity dimension; it merely implies that the hypothesis of *no change* along that dimension cannot be rejected by the data.

TABLE 2

EXTINCTION-GROUP RESPONSES IN 12 MINUTES DURING TESTS WITH GENERALIZED STIMULI

BIRD	SPOT SIZE *						
	−3	−2	−1	0	+1	+2	+3
Group I							
5	12.5 (.029)	4.5 (.012)	6.0 (.014)		13.5 (.031)	6.0 (.014)	12.5 (.029)
6	6.5 (.017)	3.5 (.009)	4.0 (.011)		1.5 (.001)	36.0 (.096)	3.5 (.009)
7	1.0 (.003)	3.0 (.009)	0.5 (.002)		3.5 (.011)	1.5 (.005)	38.0 (.111)
8	0.0 (.000)	2.5 (.007)	0.5 (.002)		0.0 (.000)	5.5 (.015)	14.5 (.039)
Group II							
21	4.0 (.004)	3.0 (.003)	2.0 (.002)	7.5 (.008)	3.0 (.003)	1.5 (.002)	17.0 (.019)
22	0.5 (.001)	1.0 (.003)	0.5 (.001)	1.5 (.004)	4.0 (.011)	1.5 (.002)	0.0 (.000)
23	1.0 (.002)	0.5 (.001)	7.0 (.018)	0.0 (.000)	4.5 (.011)	7.5 (.019)	8.5 (.021)
24	0.0 (.000)	9.0 (.017)	3.5 (.006)	4.5 (.008)	0.5 (.001)	7.0 (.013)	4.5 (.008)
Mean	3.2 (.007)	3.4 (.008)	3.0 (.007)	3.4 (.005)	3.8 (.009)	8.3 (.021)	12.3 (.030)

* The numbers in parentheses are ratios of test responses to mean responses per 12 min. in conditioning.

In the raw data it will be noted that one bird showed no reversals, one showed one, and the third gave three. Over all, the data resemble a generalization curve. The curve is asymmetrical (as in the case of the Control Group), being low on the smaller spot sizes and relatively flat beyond the standard spot size (0).

The ratio data of Table 3 are somewhat more readily interpreted than the raw results because of differences in rates of responding just prior to the test. It will be recalled that the preceding rate was allowed to vary from 45 to 55 per cent. Over all birds and all tests it averaged 49.95 per cent. The reversals tend to be minimized in the ratio figures although the major one (bird No. 16, spot size +3) remains. The curve in ratio terms is still asymmetrical with a sharper drop-off in responding on the small spot sizes than on the large.

Punishment Groups

In Table 3, for the Punishment Group, mean scores of two tests are shown for each bird and spot. Mean ratios of test responses to twice the number of responses in the immediately preceding 12 minutes with spot 0 are also presented. The factor of two in those ratios was introduced because the mean rate of responding for each bird with spot 0 during testing was 50 per cent of the prepunishment rate.

An analysis of variance of the data in Table 3 was made for both the raw data and the ratios. For the raw data, the between-spots variance yielded an F of 2.5, which is significant at about the 10 per cent level, while the between-birds variance yielded an F of 25, which is significant well beyond the 1 per cent level. For the ratios, the between-spots variance yielded an F of 3.8, which is significant at the 2.5 per cent level, while the between-birds variance gave an F of 1.1, which is insignificant. The comparison of the results of these two analyses makes it clear why the ratios were used, *viz.*, differences among birds were almost eliminated.

In order to compare the mean rates for the seven spot sizes, the method developed by Tukey (11) was used. The ordinary t-test technique is not appropriate for comparing more than two means. The results indicated that no two adjacent means were significantly different at the 5 per cent level. Nevertheless, the analysis of variance demonstrated that the hypothesis that spot size had no effect on the rate of pecking after punishment must be rejected.

An alternative hypothesis is that spot size has the same effect on responding before punishment as after punishment. To test this hypothesis, the ratio data for each spot size in Table 3 (for the Punishment Group) were divided by the mean ratios of Table 1 (for the Control Group). If the hypothesis were correct and if there were no sampling errors, these divisions would all yield a result of about 0.5. An analysis of variance of the ratios so obtained yielded an F very near unity for the between-spots variance. Hence, the alternative hypothesis cannot be rejected, and so the data do not provide any evidence for a differential generalization effect before and after punishment, except, of course, for the change in the over-all level of responding.

After the test series, the three 50 per cent shock birds were returned to the original nonshock, 12-minute APR schedule in order to study the course of recovery from prolonged punishment. In three such sessions all birds showed a recovery to their original rate early in the period followed by a gradual or abrupt decline, levelling off at about 50 per cent for two birds and 80 per cent for the third.

When response strength was reduced to about 3 per cent of its nonshock value for two additional birds, the level of responding to test stimuli was low: 5 per cent or less of original response strength except for spot size +3, where it was 15 per cent for one bird and 12 per cent for the other. The trends were the same for both birds and thus the data were averaged. The ratios of responses to preceding rate, normalized by multiplying by 3 per cent, for the spot sizes −3 through 0 to +3, were: .02, .01, .03, .02, .02, .08, and .50, respectively. The curve is low and flat through spot size +1. For the two largest spot sizes there is again an appreciable increment in responding. While the ratio of .50 looks large, the actual numbers of responses to the largest spot size were 240 and 123 as compared with nonshock rates of 1654 and 1024, respectively.

The same two birds were tested with each

TABLE 3

PUNISHMENT-GROUP RESPONSES IN 12 MINUTES DURING TESTS WITH GENERALIZED STIMULI

BIRD	SPOT SIZE *						
	−3	−2	−1	0	+1	+2	+3
13	4 (.015)	50 (.211)	98 (.385)	164 (.673)	109 (.406)	82 (.356)	70 (.273)
16	74 (.105)	240 (.310)	354 (.468)	382 (.510)	460 (.583)	270 (.345)	563 (.678)
18	156 (.278)	320 (.421)	386 (.435)	420 (.537)	323 (.388)	344 (.397)	242 (.269)
Mean	78 (.133)	203 (.314)	279 (.428)	322 (.574)	297 (.459)	232 (.366)	292 (.408)

* The numbers in parentheses are ratios of test responses to twice the number of responses in the immediately preceding 12 min. with the training spot (0).

of the seven stimuli after a criterion of 10 minutes of nonresponse had been reached. The average ratios for the seven spot sizes, where test responses were divided by the nonshock rate, were as follows: .000, .000, .001, .002, .000, .005, and .008. The largest number of responses to any stimulus for either bird was 20 (1.5 per cent of original rate). The most striking aspect of these data is their over-all low level—even below that of the extinction groups carried to the same criterion of nonresponse.

It would seem that the same interpretation applies to these last two sets of data as to the extinction results. The over-all level of responding is low and there is no significant support for an upward trend not confounded with previous training.

Three Major Groups Compared

Figure 1 summarizes the mean ratio-figures for the three major groups. The generalization decrement for the Control and Punishment Groups is striking. The low, flat curve

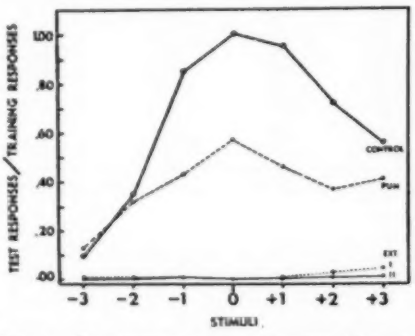

FIG. 1. GENERALIZATION AFTER CONTROL, PUNISHMENT, AND EXTINCTION CONDITIONS

for the Extinction Groups is also apparent. The separation of the three curves follows from the methodology in which the Control Group was tested when responding at a 100 per cent level, the Punishment Group when their rate was depressed to 50 per cent by shock, and the Extinction Groups when they reached the criterion of 10 minutes of nonresponse.

DISCUSSION

The results clearly show that generalization was operating in the situation studied before extinction and without punishment. Variations in rate along the dissimilarity dimension after extinction did not show significant increases in response strength. Depression of the rate to 50 per cent of the original rate by use of shock yielded essentially a generalization curve along the similarity dimension. Hence, the present data do not supply evidence in support of the Miller (7) displacement theory. Although the results of the Extinction Groups (and the 3 per cent punishment birds) are not inconsistent with a small increase in response strength, the over-all low level of responding excludes the possibility of a large effect.

The difference between the present results and those of Miller and Kraeling (8) requires comment. Those investigators found an *increase* in response strength when different runways were used after the nongeneralized response strength had been reduced to almost zero by shock. There are a number of differences between that experiment and the present ones. For example, rats were used instead of pigeons, trials were imposed instead of allowing free responding, the method of changing the stimulus situation was quite different, a running response was used instead of pecking, and 100 per cent reinforcement was used instead of partial reinforcement.[3] In the opinion of the authors, these differences are not relevant to the theoretical arguments, presented at the beginning of this paper, which made differential predictions about the outcomes of these experiments. Another difference is the following: Shock was administered through the food shield in the Miller-Kraeling study and so shock was closely correlated with eating; such a correlation was avoided in the present study.

In comparing the Miller-Kraeling experiment with the Extinction Groups of the present study, the major difference is that extinction was used instead of electric shock to reduce the response strength to near zero. It is not clear that extinction produces conflict of the sort produced by punishment, but

[3] With regard to this difference in reinforcement schedule, it is well established that partial and 100 per cent reinforcement have grossly different behavioral consequences (6). It is, however, by no means obvious how these consequences could account for the difference in results of the two sets of experiments.

one could certainly argue that extinction increases the strength of responses incompatible with the behavior originally learned.

One further point is relevant in comparing the two experiments, *viz.*, the measures of response strength used are not directly comparable. In the Miller-Kraeling study, the measure used was the number of rats that completed the response in 60 seconds on one or more of four test trials; it is not at all obvious how this measure can be related to the prepunishment response strength, and still less obvious how these measures are related to the ratios of rates of responding reported in the present investigation. It is apparent that the measures used in the two studies can be related only through a model.

A very simple model is the one used by Estes (4) and by Bush and Mosteller (1). Time is broken up into small intervals of length h, and p is the probability that an experimentally defined response will occur in such an interval. If this probability is constant during k trials each of length nh, where n is an integer, then the probability, P, that the response will occur at least once during these trials is easily shown to be given by $P=1-(1-p)^{kn}$. In an experiment, such as that of Miller-Kraeling, one should expect to estimate this quantity by the proportion of animals that completed the response during one or more of the k test trials. Knowing that proportion, and the numbers k and n, one can then estimate the probability p. It was reported that $k=4$ trials, $nh=60$ seconds, and h was about 2 seconds. The proportions 0.23, 0.37, 0.70, and 0.15 found by Miller and Kraeling then yield probabilities, p, of 0.0022, 0.0038, 0.0100, and 0.0014, respectively.

For rate problems, both the Estes model (4) and the Bush-Mosteller model (1) take the rate of responding as proportional to the probability, p. The ratios of rates shown in Table 2 for the Extinction Groups are then estimates of p; the mean ratios vary from 0.005 to 0.030 for the seven stimuli. One may note that these estimates of p are in the same range as the above estimates made from the Miller-Kraeling data. Similar ratios were obtained for the two birds tested after rate of responding was reduced by shock to near zero.

In conclusion, one can say that the post-punishment response strengths (as defined above) observed by Miller and Kraeling are of the same order of magnitude as the response strengths observed for the present Extinction Groups and for the two birds shocked to near zero. In both studies, these response strengths are very small compared to the prepunishment or pre-extinction response strengths.

From the point of view of the theories discussed above, the most clear-cut contradiction between the two sets of data comes from a comparison of the Miller-Kraeling study and our 50 per cent punishment group. We observed a significant decrease whereas they found a significant increase. Two suggestions may be offered to account for this discrepancy: (*a*) Both sets of experiments show that the "displacement increment" is small in absolute magnitude. It is further clear that the magnitude of the generalization decrement is appreciably larger when the nongeneralized response strength is 50 per cent rather than near zero. Then, if one assumes that a small "displacement increment" is operating in the 50 per cent case, it would follow that the net effect would be a decrease. (*b*) In the original stimulus situation, it may be that different aspects of that situation acquire differential degrees of importance for the two conflicting responses. For example, in our study, the illuminated window may be of major importance for the pecking response, while the cage floor may be of prime importance for the responses evoked by shock. If this is true, one might expect the choice of stimulus dimension for the test situation to determine whether a net increase or decrease would result.[4]

Even if the displacement effect were a small one or a less general one than usually supposed, it might be of considerable interest in the analysis of clinical and cultural data. The clinical evidence, however, that gave rise to the concept of displacement might be interpreted in terms of discrimination learning, i.e., in terms of differential rewards and punishments along a dissimilarity continuum (7). Such a possibility requires investigation.

[4] An experiment in which the floor texture is changed is now in process to test this hypothesis.

SUMMARY

Generalization of an instrumental response after conditioning, extinction, and punishment was studied. Three groups of four pigeons each, one group of three pigeons, and one of two were conditioned to peck an illuminated window using an irregular time schedule of partial reinforcement. The mean time between feedings was 12 minutes. The Control Group was given interspersed 12-minute, nonreward tests with three smaller light spots and three larger light spots. A clear-cut generalization curve was obtained.

Two groups of four birds each were tested with the generalized stimuli after extinction. The over-all level of responding to all stimuli was low and no significant changes in rate of responding were obtained as the stimulus was varied. The same findings resulted when two birds were tested after (*a*) shock was used to reduce response strength to 3 per cent of its original level, and (*b*) a criterion of 10 minutes of nonresponse was reached.

Another group of three birds was given additional training with food and electric shock until responding stabilized at about 50 per cent of the preshock level. Tests were then made with the various stimuli. Generalization functions were obtained; the highest rate of responding was with the training stimulus and the shape of the curve was essentially the same as that obtained with the Control Group.

The relations of these results to Miller's theory of displacement (7) and to the experiment of Miller and Kraeling (8) are discussed.

REFERENCES

1. BUSH, R. R., & MOSTELLER, F. A mathematical model for simple learning. *Psychol. Rev.*, 1951, **58**, 313–323.
2. BUSH, R. R., & MOSTELLER, F. A model for stimulus generalization and discrimination. *Psychol. Rev.*, 1951, **58**, 413–423.
3. ESTES, W. K. An experimental study of punishment. *Psychol. Monogr.*, 1944, **57**, No. 3 (Whole No. 263).
4. ESTES, W. K. Toward a statistical theory of learning. *Psychol. Rev.*, 1950, **57**, 94–107.
5. JENKINS, W. O., & CLAYTON, FRANCES L. Rate of responding and amount of reinforcement. *J. comp. physiol. Psychol.*, 1949, **42**, 174–181.
6. JENKINS, W. O., & STANLEY, J. C., JR. Partial reinforcement: A review and critique. *Psychol. Bull.*, 1950, **47**, 193–234.
7. MILLER, N. E. Theory and experiment relating psychoanalytic displacement to stimulus-response generalization. *J. abnorm. soc. Psychol.*, 1948, **43**, 155–178.
8. MILLER, N. E., & KRAELING, DORIS. Displacement: Greater generalization of approach than avoidance in a generalized approach avoidance conflict. Paper read at East. Psychol. Ass., Brooklyn, March, 1951.
9. RAZRAN, G. Stimulus generalization of conditioned responses. *Psychol. Bull.*, 1949, **46**, 337–365.
10. SKINNER, B. F. Are theories of learning necessary? *Psychol. Rev.*, 1950, **57**, 193–216.
11. TUKEY, J. W. Comparison of individual means in the analysis of variance. *Biometrics*, 1949, **5**, 99–114.

Received November 5, 1951.

RAO'S MULTIPLE DISCRIMINANT TECHNIQUE APPLIED TO THREE TAT VARIABLES [1]

HAROLD WEBSTER
University of Kentucky

PROJECTIVE techniques are most useful for investigating personality when the many variables present on a given protocol are treated as an interrelated group. It is well known that the meaning of a given predictor, indicator, or "sign" is dependent upon its relationship to other predictors in the protocol data. Statistically, this means that the intercorrelations of the predicting variables must be taken into account when it is desired to predict some criterion. This is the principle of multiple regression prediction, and any formulation of a pattern analysis for purposes of diagnosis, prognosis, or assessment should be logically reducible either to this principle or to some approximation of it.

A multiple regression method may fail to improve prediction enough to be worth the extra computation (7, 12, 29). In some cases, however, particularly with psychological data, it is not readily apparent how much the prediction is improved. For example, an investigator may have found 15 personality inventory items which discriminate a two-group criterion defined by some personality variable such as "dominance-submission." It is not usually realized that even though the items, if allowed to stand with the crude weights (erroneously called "unit" weights [3]), may misclassify *no more* members of the sample than do the same items with linear multiple regression weights ("best" weights), those misclassified may nevertheless include *different* individuals in the two cases. The consequences of this are crucial: multiple regression weights are always preferable to other weights, for they use the maximum information available in the scores (or in functions of the scores). Investigators who compute regression weights should apply the tests for the precision of discrimination which the resulting functions achieve (6, 11, 20).

Multiple regression weights become less useful as the number of items is increased (30). Bedell (1) has shown, however, that when the criterion is obscure the reliability of a group of items may decrease as the items become more numerous. There is also other evidence that a small number of items may do efficiently the work of many (14, 28). It seems likely that much projective data can be valuable for predictive purposes if scored on relatively few variables (or "items"), provided the latter may be used in the appropriate regression functions.

PROBLEM

The purpose of this paper is to describe an application of the relatively new method of multiple regression pattern analysis developed by Rao (19, 20) to Thematic Apperception Test (TAT) scoring variables.

The problem was complicated, as was expected, by the fact that TAT scores of interest are obtainable only by rather inexact measurement procedures. It may be assumed that most projective protocol variables, like other personality variables (2, 4, 23, 25, 26), are *continuously* distributed in the population. Two serious difficulties arise, however, in computing regression functions for projective data. The first is that the obtainable scores for variables of importance are often so categorical, or "qualitative," that the presence of non-normal distributions and nonlinear regressions may go undetected. The second difficulty is that an impractically large amount of computation may be necessary. For example, when data are known to require nonlinear functions, the computation is excessive (16, 27).

There is no solution to the problem of using discrete attributes, or purely qualitative categories, in multiple regression prediction (16). Consequently, it must be decided either to scale every variable so that it will be truly quantitative (10, 14, 15), or to assume that "qualitative" scores are degenerate quantitative measures which may nevertheless be used

[1] This study was submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy at the University of California at Berkeley. The author wishes to acknowledge the assistance of Professors D. W. MacKinnon, R. N. Sanford, R. F. Jarrett, and D. H. Lehmer of the University of California.

in such a way that the principle of multiple regression prediction is approximately satisfied. It is preferable to make the latter assumption only because the methods of obtaining true scales are still relatively complicated and undeveloped. A method which in practice would make such an assumption has been suggested by Guttman (9, p. 274). Thus variates which can only be said, with high interscorer agreement, to be either present or absent would be scored either unity or zero. If a variate were detectable only in low, medium, or high amounts, then the corresponding scores would be $-1, 0, +1$ or $0, +1, +2$.

For projective data, it may remain unknown how closely such restricted scores represent the continuously distributed underlying variates. It would appear from studies with similar types of psychological data, however, that there is little to be gained from more precise scoring (28, especially p. 547). A small set of constants used for scoring also has the distinct advantage of greatly simplifying computations. Since it is known that scores are automatically standardized by an operation required in deriving the regression functions, namely, the inversion of the covariance matrix (24), the effects of the inexactness of such scores are more apparent than real. In some cases where more exact measurement is feasible, and where only small samples may be used, it may not be advisable to use a few constants as scores. In recent studies, however, extensive use has been made of them (8, 17, 23, 28, 31).

Rao's method, which is essentially a generalization of Fisher's (5) discriminant function technique for two groups, is superior to other methods that might be used for pattern analyses of projective data because it may assign individuals to *more than two* criterion groups. It is the only method by which it is possible to assign, *with minimum errors of classification*, N people to q criterion groups on the basis of all information in k intercorrelated measures.

PROCEDURE

Subjects

The subjects were 315 freshman and sophomore college men from speech and history classes at the University of California at Berkeley. The intercorrelated projective measures were three scoring variables from two TAT stories. The criterion groups, to be defined below, were four subgroups of the sample.

How the Data Were Obtained

The two TAT pictures and a miscellaneous group of questionnaire items to be used for defining criterion groups were mimeographed and group-administered. The TAT pictures were versions of the mother-son card (6BM) and the hypnotist card (12M), both from the Murray series (18). They were administered in this order before the questionnaire items were answered.

Murray's description of the first card is, "A short elderly woman stands with her back turned to a tall young man. The latter is looking downward with a perplexed expression." His description of the second card is, "A young man is lying on a couch with his eyes closed. Leaning over him is the gaunt form of an elderly man, his hand stretched out above the face of the reclining figure." Although these main features of the pictures were preserved in the mimeographing process, many of the smaller details were not.

The *Ss* wrote their own stories under the mimeographed pictures during the first 20 minutes of the class period and used the remaining time to check the blanks in the part of the form consisting of questionnaire items. About 4 per cent of the forms were unsuitable as data because of skipped pages or unusual facetiousness or hostility in the stories.

After the TAT stories had been scored the entire scores for each form were punched into cards so they could be mechanically counted and sorted.

Definition of the Four Criterion Groups

The aim in defining criterion groups was twofold. It was desired to define groups which could be said to differ in personality characteristics, and it was believed that characteristics which would be likely to be associated with differences in TAT stories should be used. It was supposed that characteristics fulfilling these requirements might be the scores on the two traits:

a. Expressed preference for being submissive to authority and convention.

b. Expressed symptoms of emotional insecurity and nonsocial feeling.

Neither of these traits is known to be very homogeneous in the sense of containing only one psychological dimension, but both are important to clinicians and would also be expected to be associated with differences in TAT stories. Accordingly, the following six "submission-dominance" items were allowed one point each,[2] when answered as indicated, for a "submission" score:

1. Would rather work for a "worthy superior" than for self.
2. Attends church "frequently."
3. "Usually agreed" with father on social and political affairs.

[2] Best weights would have been used if an external criterion had been readily available.

4. "Usually agreed" with mother on social and political affairs.

5. Favored the authoritative parent (based on 2 items).

6. "Best" at "mechanical things" or "artistic and literary things" rather than "logic and argument."

The following eleven "secure-insecure" items were allowed one point each,[2] unless otherwise indicated, for an "insecurity" score:

1. Prefers not to work mainly with people.

2. Omits the marriage question or "never" wants to marry.

3. Omits the children question or wants only one child or no children.

4. Drinks "frequently."

5. Participates in no extracurricular activities.

6. Omits answer or writes in "Nothing" to "At what things are you best?"

7. Omits answer to "What things or situations are you most afraid of?".

8. Lists only phobic, sex, religious, or health fears (scored 0, 1, or 2).

9. Expresses high dislike of people's weaknesses (based on 28 items).

10. Expresses high dislike of sex expression in others (based on 4 items).

11. Expresses high dislike of sloppiness or dirtiness in others (based on 2 items).

When cuts were taken at the means of both these distributions, the sample was "quartered" in such a way that the criterion groups could then be designated, briefly, as submissive-secure, submissive-insecure, dominant-secure and dominant-insecure. Abbreviations followed by frequencies for these groups comprise the first columns of Table 1.

The validity or ultimate empirical usefulness of the criterion groups is, of course, open to question. Ideally such groups would be defined only after intensive investigation of the traits in each individual. In much personality research, however, the individual investigator is restricted, for practical reasons, to less intensive methods. Presumably the criterion groups of the present study are defined adequately enough to be of some interest to psychologists and, at the same time, to serve as an example of how better-defined groups may be utilized in subsequent work with projective data.

The TAT Variables

After experimenting with a number of scoring systems, 20 TAT scoring variables (ten for each of the two stories) were finally found which gave agreement of $r > 0.82$ for two scorers working independently. Both TAT scores and questionnaire item responses had been punched into cards, and it was therefore possible to machine-sort the 20 TAT variables to reveal differences among the four criterion groups described in the previous section. Three scores selected as potential discriminators of the four groups were:

A. In the first TAT story (to the mother-son picture) the man tries to do what is best either for himself or for someone else not in the picture. That is, he is not primarily interested in the welfare of the woman, nor is he interested in everyone's or no one's welfare (neutral category).

B. In the first story the woman tries to do what is best either for herself or for someone else not in the picture. That is, she is not primarily interested in the welfare of the man (or in everyone's or no one's welfare).

C. In the second story (to the hypnotist picture) the reclining man is described as voluntarily hypnotized.

When one of these three conditions applied for a given story, the *S* received a score of 1 on that variable; *Ss* for whom the condition did not apply received a score of 0.

The Statistical Analysis

Further statistical details relevant to the procedure to be described may be found in the articles by Rao (19, 20, 21, 22) and in the paper by Rao and Slater (23). In the latter the method is used with psychological diagnostic data.

In Table 1 the numbers of cases in each group, for each of the three TAT scores, are followed by the corresponding means. For example, 39 of the 87 men in the submissive-secure group used the TAT variable as it is described in B above.

TABLE 1

FREQUENCIES AND MEANS FOR THREE TAT VARIABLES FROM STORIES BY FOUR GROUPS OF MEN

GROUP		TAT VARIABLES*					
		A		B		C	
	n	*f*	*M*	*f*	*M*	*f*	*M*
Sub–Ins	60	23	.383	15	.250	13	.217
Sub–Sec	87	51	.586	39	.448	33	.379
Dom–Ins	89	44	.494	28	.315	26	.292
Dom–Sec	79	28	.354	21	.266	10	.127
Total	315	146		103		82	

* The abbreviations for the groups and for the TAT variables A, B, and C are explained in the text. In this table and in the subsequent tables, the figures have been rounded from four or more decimal places to only three or four.

In examining Table 1 we may ask whether the three intercorrelated TAT variables give better than chance discrimination of the four groups. In order to answer this question, an analysis of dispersion, or generalized analysis of variance (20), was performed as illustrated in Tables 2, 3, and 4.

TABLE 2

TOTAL PRODUCT-SUM MATRIX WITH 314 *df*

	A	B	C
A	78.330	23.108	—7.127
B	23.108	69.321	0.783
C	—7.127	0.783	60.654

Elements of Table 2 are sums of squares and products of the three TAT variables, or $a_{ij}-N\bar{x}_i\bar{x}_j$, where a_{ij} is the sum of joint occurrences of scores of unity for variables i and j, N is the total number of cases, or 315, and the $\bar{x}_i$ and $\bar{x}_j$ are the means of the two variables. The a_{ij} were easily obtained in a few minutes by sorting the punched cards, and the means were obtained from Table 1. This expression for elements of the total within-groups dispersion matrix is a special case, for two-valued variates, of the general expression for covariance. It may be seen that it is N times the covariance of a fourfold table.[3] When divided by N, elements of this form are the same as those used by Lazarsfeld (13, pp. 369–379) in "latent structure analysis," a method by which manifest attributes are analyzed in order to infer the characteristics of latent continua postulated to produce them.

TABLE 3

PRODUCT-SUM MATRIX BETWEEN GROUPS WITH 3 *df*

	A	B	C
A	2.720	2.158	2.720
B	2.158	1.945	2.068
C	2.720	2.068	2.102

Table 3 is obtained directly from Table 1 and has 3 *df*, or one less than the number of groups. Its elements represent portions of the elements of Table 2 attributable to differences between groups. When it is subtracted from Table 2, Table 4 is obtained.

TABLE 4

PRODUCT-SUM MATRIX WITHIN GROUPS WITH 311 *df*

	A	B	C
A	75.610	20.950	−9.847
B	20.950	67.376	−1.285
C	−9.847	−1.285	58.552

Table 4 has $N-q$, or 311 *df*, one *df* having been lost for each of the q groups.

To test for the discrimination of the four groups by the three TAT variables, chi square is computed (20, pp. 67–71) as

$$\chi^2=-(N-1-\tfrac{1}{2}[k+m+1])\log_e d,\quad df=km,$$

where N=total number of cases, or 315
k=number of measurement variables, or 3
m=number of groups minus one, or $q-1=3$
d=the ratio of the determinant of Table 4 to the determinant of Table 2
e=the base for natural logarithms.

For the present problem d was found to be 0.90884 so chi square with 9 *df* is 29.681. This value has a small probability, $p<.001$, of occurring by chance, which probably does indicate some discrimination in this sample. Because of the selection from the 20 TAT variables of three which seemed most significant, however, the test must be viewed with caution.

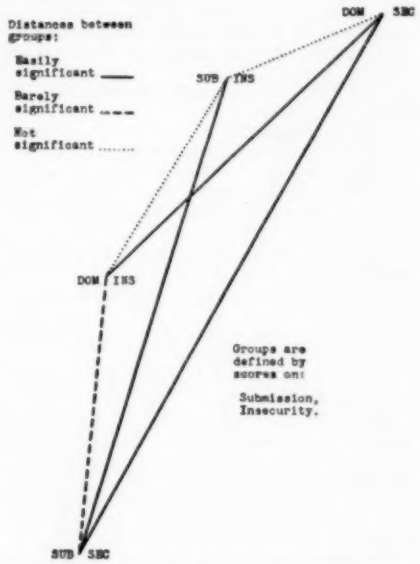
Distances between groups:
Easily significant
Barely significant
Not significant
DOM
SUB
INS
SUB
INS
DOM
SUB
SUB
INS
DOM
SUB
INS
DOM
SUB
SUB
INS
DOM
SUB
INS
DOM
SUB
SUB
INS
DOM

TABLE 6

VALUES OF D^2 ARRANGED FOR EACH GROUP WITH EVERY OTHER GROUP IN ASCENDING ORDER OF MAGNITUDE

SUB-INS			SUB-SEC
Dom-Sec	0.0548	Dom-Ins	0.1448
Dom-Ins	0.0999	Sub-Ins	0.4595
Sub-Sec	0.4595	Dom-Sec	0.7193
DOM-INS			**DOM-SEC**
Sub-Ins	0.0999	Sub-Ins	0.0548
Sub-Sec	0.1448	Dom-Ins	0.2646
Dom-Sec	0.2646	Sub-Sec	0.7193

Using elements of Table 5, the D^2 were obtained and are given in ascending order for each group in Table 6. The D^2 for each group were tested for significance (11), by

$$F=n_1n_2(n-k+1)D^2/(n_1+n_2)nk,$$
$$k \text{ and } n-k+1 \; df,$$

where k is the number of variates, n is the df for the c^{ij} matrix and n_1 and n_2 are numbers in the two groups. Table 7 gives the values for these F-ratios, which correspond to the distances between group means as they are shown in Figure 1. (Psychological interpretations of Tables 1 and 7 and of Fig. 1 are given under *Results*.)

TABLE 7

VARIANCE RATIOS FOR VALUES OF D^2 IN TABLE 6 *

	SUB-SEC	DOM-INS	DOM-SEC
Sub-Ins	4.162	1.185	0.619
Sub-Sec		2.110	9.863
Dom-Ins			3.668

* The value of the ratio, for 3 and 309 df, is 2.63 at the .05 level of confidence and 3.85 at the .01 level.

It is possible to fit a line, plane, or higher element to points representing group means. In the present case these groups actually have three coordinates corresponding to the three TAT variables, but the residual dispersion from the *plane* of Figure 1 was so slight that it was not tested for significance.[4] The procedure for fitting means in a multivariate space (19) is a completely unambiguous method of factor analysis known as "canonical" analysis. Figure 1 shows that at most two factors, corresponding to the two dimensions of the plane, are sufficient to account for the data.

The four groups are defined by the two sets of items used as criteria but the distances between them are due to the three TAT variables when allowance has been made for the intercorrelations among the latter. This is important for it means that both the TAT predictors and the six distances between the group means are patterns determined by all information in the individual scores.

Relatively little information is contributed by only three TAT variables, each of which can have one of only two possible values. For each group there are only eight (2^3) possible patterns of scores, but the analysis is the same as if there had been more. For example, if the TAT variables had been represented by five-point scales, there would have been 125 (or 5^3) possible patterns of scores for each group. The method is general and can be used [5] with any number of mutually exclusive criterion groups, as well as with any number of TAT scores from any size scales. Even a pattern analysis of only three TAT "signs" reveals information, however, that would not have been discovered if the three signs had been used singly as predictors.

A canonical factor analysis is not necessary in order to compute the discriminant functions. The latter are used for assigning any individual, on the basis of his TAT scores, to the group of which he is most likely to be a member. Table 8 gives the coefficients k_{gi} and the constant terms K_g for the linear discriminants

$$L_g=k_{g1}x_1+k_{g2}x_2+k_{g3}x_3+K_g, \; g=1, \ldots, 4,$$

for the four groups.[6] The x_1, x_2, x_3 are 1 or 0,

[4] True proportions for distances between groups in Figure 1, starting at the top corner with the Dom-Sec group and proceeding clockwise around the perimeter, are 8.48, 3.80, 3.16, and 2.34, the two inner distances being proportional to 6.78 and 5.14.

[5] The method is facilitated by using orthogonal transformations of the original scores to avoid direct inversions of the matrices when the latter must be large (21).

[6] $K_g=\log_e P_g-\frac{1}{2}(k_{g1}\bar{x}_{g1}+k_{g2}\bar{x}_{g2}+k_{g3}\bar{x}_{g3})$ where $k_{g1}=c^{11}\bar{x}_{g1}+c^{12}\bar{x}_{g2}+c^{13}\bar{x}_{g3}$, $k_{g2}=c^{21}\bar{x}_{g1}+c^{22}\bar{x}_{g2}+c^{23}\bar{x}_{g3}$ and $k_{g3}=c^{31}\bar{x}_{g1}+c^{32}\bar{x}_{g2}+c^{33}\bar{x}_{g3}$. The c^{ij} are elements of the inverse dispersion matrix of Table 5 and the $\bar{x}_{gi}$ are means obtainable from Table 1.

TABLE 8

VALUES NEEDED TO COMPUTE THE LINEAR DISCRIMINANTS L_g FOR THE FOUR GROUPS *

GROUP	COEFFICIENTS				CONSTANTS, K_g	
	L_g	k_{g1}	k_{g2}	k_{g3}	GENERAL PROPORTION	FOR THIS SAMPLE
Sub-Ins	L_1	1.576	0.692	1.433	$\log_e P_1-0.544$	−2.202
Sub-Sec	L_2	2.352	1.387	2.444	$\log_e P_2-1.464$	−2.750
Dom-Ins	L_3	2.053	0.852	1.918	$\log_e P_3-0.922$	−2.186
Dom-Sec	L_4	1.353	0.825	0.919	$\log_e P_4-0.408$	−1.791

* Use of discriminants, $L_g=k_{g1}x_1+k_{g2}x_2+k_{g3}x_3+K_g$, is discussed in the text.

depending upon whether the corresponding TAT scores (A, B, C) are present or absent. The K_i in the last column of Table 8 are given for proportions of the four groups of the present study. For other samples, the P_i must be recomputed. The use of the functions with other samples, however, will introduce an unknown amount of sampling error.

The procedure in using the discriminants is to compute L for each group and then assign the individual to the group for which it is greatest.

The analysis which has been described is adequate for treating several independent and several dependent personality or projective variables simultaneously. An assumption in addition to the assumption of continuity of the population variates became necessary when the tests for significance were made. The tests assumed normal distributions in the population.

RESULTS

The means in Table 1 reveal that proportionally fewer members of all four groups used B than used A, and that even fewer used C than used B. It seems reasonable that these TAT scores would show this variation for young men. Young persons of both sexes[7] identify with an autonomous young man in the first picture more often than with an autonomous old woman who does not care for him. Also, the men were not likely to see the reclining figure in the second picture as voluntarily submitting to the older man, especially if they were in other than the submissive-secure group.

It can be seen from the score patterns of Table 1 that the sub-ins and dom-sec groups may be more alike than any other pair of groups. This is confirmed by Table 7 and Figure 1.

When Table 7 and its geometrical picture, Fig. 1, are examined, several facts of psychological importance emerge, facts which would not have been apparent if the TAT variables had been used singly as predictors.

First, within the definitions of "security," "submission," etc., used in this study, the most effective discrimination was discrimination of dominant men from submissive men among the secure group (note the longest axis in Fig. 1). In fact, among the insecure, discrimination of dominant from submissive, although not singly significant at the 5 per cent level, was inverted when compared with the secure group. Therefore, the three TAT scores for dominant and submissive men differed greatly only among the secure group, and if they differed at all among the insecure it was not in the same way. This is evidence, as revealed by TAT scores, that the insecure do not discriminate between dominant and submissive roles as well as the secure, and that when they do they are likely to invert the roles.

Second, among both dominant and submissive men (note the two crossed axes in Fig. 1), TAT stories discriminated the secure from the insecure. Putting this finding with the first above, situations may now be visualized, with the help of Fig. 1, in which the type of sample may determine whether or not differences will be obtained on a variable of interest in TAT stories.

For example, if we should choose men who were both dominant and insecure, we would expect their stories to differ on the three variables from those of dominant men who were also secure *more* than they would differ from the stories of submissive men (whether secure or insecure). Or, if we should sample submissive men, we should predict that those among them who were insecure would write stories which would differ more from the stories of those who were secure than they would differ from the stories of dominant men (whether secure or insecure).

It can be seen how preselection on even one personality variable can affect TAT stories, or, conversely, how the predictive value of TAT stories may be drastically changed by one personality variable. Note, for example, that although there are two axes in Fig. 1 which clearly discriminate secure from insecure, and one which definitely discriminates dominant from submissive, the dominant-secure group is nevertheless insignificantly discriminated from the submissive-insecure group.

From the evidence just presented it may be seen that TAT stories are obviously sensitive to personality variables, but that they are for that reason extremely complicated prediction devices for which single-variable prediction procedures may be inadequate. In other words, a single correlation coefficient suitable for use in a regression equation containing only one independent TAT variable may be so sensitive to extraneous personality variables that it is not only useless but also misleading.

[7] A comparable analysis for 184 college women resulted in the means (analogous to those for the men in Table 1), reading by rows from left to right: .525, .400, .300; .500, .333, .315; .429, .265, .286; and .463, .342, .342. These means gave insignificant discrimination of the four groups of women but are nevertheless of interest here for comparison with the men's means.

A further interpretation of the results is possible from a study of the discriminant functions themselves and Table 9. In Table 9 are given the values of L_g for the eight possible score patterns.

In the following paragraph the ones and zeros in parentheses refer, in order, to the presence or absence of the three TAT scores A, B, and C. The latter give values of unity or zero, for individuals to be classified, to the x_1, x_2, x_3 for use with Table 8.

Men closest to the dom-sec group mean (Fig. 1) are those who used none of the scores (0, 0, 0) or who saw the woman as not helping the man (0, 1, 0). The latter are, however, more secure (though no less dominant) than the former. Men closest to the dom-ins mean are those who saw a man who was not interested in the woman's welfare (1, 0, 0) or a man voluntarily requiring attention from another man (0, 0, 1), those scoring (0, 0, 1) being more submissive and secure than those with scores (1, 0, 0). Those men who saw both one of the men and the woman as not actively helping the other character (1, 1, 0 or 0, 1, 1), or who saw both the men and the woman in such roles (1, 1, 1), are relatively submissive and secure. Those who saw both a man not helping the woman and a man voluntarily submitting to another man (1, 0, 1) are also closest to the sub-sec mean but are nevertheless more insecure than other groups who used more than one of the scores.

It can be seen from the preceding paragraph that men who can allow the woman, or mother figure, to be disinterested in the man's welfare are relatively secure. In fact, an interpretation of the roles given to the men in both pictures is largely dependent upon the presence or absence in the first story of such an autonomous role for the woman. Aside from its intrinsic interest, this constitutes another way of demonstrating that it is necessary to treat signs in projective data as patterns rather than as single prognostic indicators.

TABLE 9

L_g VALUES OBTAINED WITH THE EIGHT POSSIBLE SCORE PATTERNS

x_1	x_2	x_3	SUB–INS, L_1	SUB–SEC, L_2	DOM–INS, L_3	DOM–SEC, L_4
0	0	0	−2.202	−2.750	−2.186	−1.791
1	0	0	−0.626	−0.398	−0.133	−0.438
0	1	0	−1.510	−1.363	−1.334	−0.966
0	0	1	−0.769	−0.306	−0.268	−0.872
1	1	0	0.066	0.989	0.719	0.387
1	0	1	0.807	2.046	1.785	0.481
0	1	1	−0.077	1.081	0.584	−0.047
1	1	1	1.499	3.433	2.637	1.306

SUMMARY AND CONCLUSIONS

The problems which arise in applying multiple regression analysis to projective data were discussed. By use of a method of Rao's, three TAT variables (reliably scored as "signs") were shown to discriminate four criterion groups in a sample of 315 men. The four groups had been defined by two sets of personality questionnaire items, one set giving a measure of submissiveness, the other a measure of insecurity.

Simplified scoring was discussed and used. The covariance matrix required by the analysis (Table 2) was obtained very easily by mechanically sorting cards into which the scores had been punched.

Differences in TAT score patterns among the criterion groups were found. As a result of the analysis, several facts of psychological importance which would not otherwise have been apparent emerged:

1. The TAT variables significantly discriminated dominant from submissive men only among those who were secure.

2. As interpreted from TAT scores, the insecure *S*s did not discriminate between dominant and submissive roles as well as the secure, and when they did they were likely to invert the roles.

3. Among either dominant or submissive men, the TAT variables discriminated the secure from the insecure.

4. Even though there were three axes of significant discrimination among the four groups (Fig. 1), the dominant-secure group was not discriminated from the submissive-insecure group.

5. A study of the linear discriminant functions for the four groups revealed, among

other facts, that an interpretation of the roles given to the men in the TAT stories was largely dependent upon the presence or absence in the first story of an autonomous or selfish woman.

Inasmuch as a multiple regression pattern analysis was shown to be required for both understanding and utilizing the predictive meaning of only three intercorrelated TAT "signs," it would seem that those who wish to use TAT scoring variables, or comparable projective data, for purposes of prognosis or assessment should:

a. Use single correlation coefficients for purposes of prediction (from the single independent variables) only when samples can be obtained in such a way that there is either complete control or randomization of other personality variables.

b. Use pattern analysis which approximately satisfies the principle of multiple regression prediction when the conditions in *a.* are impossible to fulfill.

REFERENCES

1. BEDELL, B. J. Determination of the optimum number of items to retain in a test measuring a single ability. *Psychometrika*, 1950, **15**, 419–430.
2. BURT, C. The analysis of temperament. *Brit. J. med. Psychol.*, 1938, **17**, 158–188.
3. BURT, C. The influence of differential weighting. *Brit. J. Psychol., Statist. Sect.*, 1950, **3**, 105–123.
4. EYSENCK, H. J. *Dimensions of personality*. London: Routledge and Kegan Paul, 1947.
5. FISHER, R. A. The use of multiple measurements in taxonomic problems. *Ann. Eugen., Camb.*, 1936, **7**, 179–188.
6. FISHER, R. A. The precision of discriminant functions. *Ann. Eugen., Camb.*, 1940, **10**, 422–429.
7. GUILFORD, J. P. *Fundamental statistics in psychology and education*. (2nd Ed.) New York: McGraw-Hill, 1950.
8. GUILFORD, J. P., & MICHAEL, W. B. *The prediction of categories from measurements*. Beverly Hills, California: Sheridan Supply, 1949.
9. GUTTMAN, L. An outline of the statistical theory of prediction. In P. Horst, P. Wallin, *et al.*, *The prediction of personal adjustment*. New York: Social Science Research Council Bulletin No. 48, 1941. Pp. 253–312.
10. GUTTMAN, L. A basis for scaling qualitative data. *Amer. sociol. Rev.*, 1944, **9**, 139–150.
11. HOTELLING, H. The generalization of Student's ratio. *Ann. math. Statist.*, 1931, **2**, 360–378.
12. KLEIN, G. S. An application of the multiple regression principle to clinical prediction. *J. gen. Psychol.*, 1948, **38**, 159–179.
13. LAZARSFELD, P. F. Latent structure analysis. In S. A. Stouffer, L. Guttman, *et al.*, *Measurement and prediction*. Princeton: Princeton Univer. Press, 1950. Chapters 10–11.
14. LOEVINGER, JANE. A systematic approach to the construction and evaluation of tests of ability. *Psychol. Monogr.*, 1947, **61**, No. 4 (Whole No. 285).
15. LOEVINGER, JANE. The technic of homogeneous tests compared with some aspects of "scale analysis" and factor analysis. *Psychol. Bull.*, 1948, **45**, 507–529.
16. LUBIN, A. Linear and nonlinear discriminating functions. *Brit. J. Psychol., Statist. Sect.*, 1950, **3**, 90–103.
17. MICHAELS, W. C., & HELSON, H. A reformulation of the Fechner law in terms of adaptation-level applied to rating-scale data. *Amer. J. Psychol.*, 1949, **62**, 355–368.
18. MURRAY, H. A. *Thematic Apperception Test manual*. Cambridge: Harvard Univer. Press, 1943.
19. RAO, C. R. Utilization of multiple measurements in problems of biological classification. *J. roy. Statist. Soc. Ser. B*, 1948, **10**, 159–203.
20. RAO, C. R. Tests of significance in multivariate analysis. *Biometrika*, 1948, **35**, 58–79.
21. RAO, C. R. On a transformation useful in multivariate computations. In P. C. Mahalanobis, D. N. Majumdar, *et al.* Anthropometric survey of the United Provinces, 1941: a statistical survey. *Sankhyā*, 1949, **9**, 89–324. Appendix 5, pp. 251–253.
22. RAO, C. R. On some problems arising out of discrimination with multiple characters. *Sankhyā*, 1950, **9**, 343–366.
23. RAO, C. R., & SLATER, P. Multivariate analysis applied to differences between neurotic groups. *Brit. J. Psychol., Statist. Sect.*, 1949, **2**, 17–29.
24. SCHMIDT, J., JR. A comparison of two procedures for calculating discriminant function coefficients. *Psychometrika*, 1950, **15**, 431–434.
25. SLATER, E. The neurotic constitution. *J. Neurol. Neurosurg. Psychiat.*, 1943, **6**, 1–16.
26. SLATER, E., & SLATER, P. A heuristic theory of neurosis. *J. Neurol. Neurosurg. Psychiat.*, 1944, **7**, 49–55.
27. SMITH, C. A. B. Some examples of discrimination. *Ann. Eugen., Camb.*, 1947, **13**, 272–282.
28. STOUFFER, S. A., GUTTMAN, L., *et al.* *Measurement and prediction*. Princeton: Princeton Univer. Press, 1950.
29. THORNDIKE, R. L. *Personnel selection*. New York: John Wiley, 1949.
30. WILKS, S. S. Weighting systems for linear functions of correlated variables when there is no dependent variable. *Psychometrika*, 1938, **3**, 23–40.
31. YATES, F. The analysis of contingency tables with groupings based on quantitative characters. *Biometrika*, 1948, **35**, 176–181.

Received November 10, 1951.

VERBAL ATTITUDES AND OVERT BEHAVIOR INVOLVING RACIAL PREJUDICE [1]

BERNARD KUTNER
The American Jewish Congress

CAROL WILKINS AND PENNY RECHTMAN YARROW
Sarah Lawrence College

THE problem of explaining discrepancies between verbal attitudes and actual behavior relevant to these attitudes has been the subject of considerable theoretical thought (1). Empirical investigations of this problem have been exceptionally sparse. The classic study of LaPiere (2) demonstrated that when an innkeeper was confronted with a white person and a Chinese couple "in the flesh," requesting lodging, the trio were almost never denied it. Yet when a mailed request was sent to the same innkeeper requesting lodging for himself and his Chinese companions there was an almost universal rejection of the request. What is the nature of the situation in each case that leads to such diametrically opposed behaviors? Put in another way our question is: Why is there a discrepancy between behaviors supposedly dependent upon the same constellation of attitudes?

Our procedure was as follows:

Three young women, 2 white and 1 Negro, all well-dressed and well-mannered, entered 11 individual restaurants in a fashionable Northeastern suburban community, Subtown. In each case, the white women entered first, asked for a table for three and were seated. The Negro woman entered a short while later, informed the hostess or headwaiter that she was with a party already seated, found her table and sat down. This procedure was repeated in each of the 11 restaurants and taverns.

In each case all three women were served in a manner in no wise different from the usual service accorded patrons at each establishment. At two places the group was taken unusual notice of by other patrons but in no instance was anything but exemplary service accorded them by the management, waiter, etc.

Two weeks following each visit, a letter signed with an assumed name was sent to each establishment. The return address was that of a cooperating resident of Subtown. The letter read as follows:

Dear Sir:

A group of friends and I are planning a social affair to be held in Subtown in the near future. I should like to make reservations to have them for dinner at your restaurant. Since some of them are colored, I wondered whether you would object to their coming.

Could you let me know if the reservations may be made so that I may complete the arrangements as soon as possible?

Yours truly,
M..... B.....

Where appropriate, the term "beer party" was substituted for "social affair" and "tavern" for "restaurant."

Since no letters were returned by the Post Office,[2] it may be assumed that all were delivered. Seventeen days after the letters were sent out, no replies of any kind had been received. At this point each establishment was called by phone by one of the white women and essential parts of the letter repeated. In each instance the manager or his representative answered the phone and he was asked for reservations for the party. One day later "control calls" were made by the same person. She merely requested reservations for a party of friends to be held in the near future. Table 1 summarizes the nature of the responses to the letter, and to each phone call.

It will be noted that no spontaneous replies to the letter were received although two letters were received as a consequence of the first phone call. Three of the 11 managers

[1] The authors feel greatly indebted to Dr. Isidor Chein and Dr. Harold Proshansky for reading and criticizing of the manuscript.

[2] As an added precaution, the local postman was requested to deliver to the address used any returned letters or regular mail bearing the fictitious name on the envelope.

admitted having received the letter, eight denied having received it.

With regard to reservations, except for Restaurant D (which has a "no reservation needed" policy), every control call terminated with a specific reservation. Ignoring Restaurant D, no manager accepted the phoned request for reservations when told that Negroes would be in the party. Three managers claimed that they do not take reservations; another said that he would accept a phoned reservation, but hung up before it could be made. Three maintained that they could make "arrangements" only if someone came in to the restaurant about it. One of the latter group suggested the possibility of a separate, private room for the dinner. Two managers suggested that we call back at another time. One of these managers went as far as to visit the address on our letter and demanded to know who the writer was and why she was causing a disturbance. During this visit he obtained the phone number and subsequently phoned back three times asking to speak to our fictitious writer.

Finally, one manager turned down the phoned request on the grounds that it was against the law to admit Negroes to his tavern.[3]

It is obvious from Table 1 that the managers were disturbed by both the letter and the first call—they wished to avoid, if possible, having any Negro patrons but they also did not (for whatever reason) wish to give a categorical "No" to the request. To avoid the possibility of a charge of discrimination, they tried to dissuade the caller from coming to their restaurant, denied that they took reservations, requested a personal interview to discuss the matter, etc.

Nevertheless, five managers finally gave tentative approval for the party, two confirming the conversation with a letter. While none of these agreements was made unconditionally, tacit agreement was implied. Three others indicated that while they are unprejudiced themselves (Negro help in kitchen, some Negroes eat there, etc.) they felt that it would be better if our assistant "listened to reason" and went elsewhere.

To call the behavior of the managers inconsistent with either their attitudes or their alleged previous behavior seems trite. The manager or owner of a business that depends on majority group patronage (which he may perceive, rightly or wrongly, as hostile to minority group "intruders") may nevertheless not object when a minority group member appears on the premises. Several hypotheses suggest themselves as possible explanations of his behavior:

1. He fears the consequences of violating the law forbidding discriminatory practices.

2. He may wish to avoid a scene that might not only be a disturbance of the peace but might also call particular attention to the presence of the unwanted guest.

3. Without consciously realizing what the rationale of his behavior may be, the reality situation moves him to adopt one of two competing motives: (*a*) to do nothing, thus preserving peace and order, or (*b*) to refuse admission or service because he or his white patrons are offended by Negroes.

Whatever the underlying explanation may be, three levels of response have been found in the present study. These are reactions to each of the three situations with which each manager was confronted. The letter, which may be described as a relatively impersonal contact, produced uniform results: no reply, or at most, one very belated reply. The phone call, which may be described as a partial personal contact, produced a variety of responses ranging from implied acceptance to attempts to rationalize or justify outright rejection of the requests with numerous intermediary responses ("call back," "come in and see us," "couldn't you go to another place?", etc.). The direct contact, which may be described as fully personal in nature, also produced uniform results: admission of the Negro and excellent service in each establishment.

Thus, in addition to the contradictory behaviors found by LaPiere, the managers in the present study exhibited relatively consistent techniques designed to dissuade or discourage our requests for reservations when such requests were made by phone.

Taking the three situations into consideration it would appear, at least for our sample, that the mode of dealing with a minority group member whose presence "violates" cul-

[3] A state law on discrimination does exist, but it *forbids* denial of service to a person in a public place on the grounds of race, religion, or creed.

TABLE 1

SUMMARY OF RESPONSES TO LETTER AND PHONE CALLS

RESTAURANT	RESPONSE TO LETTER	TELEPHONE CALL	CONTROL CALL
A	No reply	Didn't get any letter. We don't take large parties. We've got dancing after 6 P.M. [They actually don't.] Are you colored? [Yes.] I like everyone. My kitchen help are colored and they are wonderful people. But we have a certain clientele here. . . . This place is my bread and butter. Frankly I'd rather you not come. Try in T——— [next town].	Took reservation.
B	No reply	I didn't get your letter. We can't have you. It's against the law.	Took reservation.
C	No reply	I got the letter. How many? [Negroes] I don't mind but customers might. In fact some of my help are colored. I had trouble about this before. Frankly I prefer you don't come but if you can't find another place we won't embarrass you here.	Took reservation.
D	No reply	We didn't get a letter. We don't take reservations. We take care of our regulars. A few Negroes come in just to eat. I *would* mind you coming.	Said they didn't take reservations but we should come in any time about them.
E	No reply	I didn't get any letter. Call any time to make reservation. [Hung up before reservation could be made.]	Took reservation.
F	No reply	Yes, I got your letter. I refuse to discuss it on the phone. What is the purpose of your party? What school did you attend? Why are you never home when I call? How many Negroes are in the party? [5 out of 10] That's a large percentage isn't it? [Insisted he call us, instead of vice-versa.]	Took reservation.
G	Reply received 19 days after our letter was mailed.	A letter to you is in the mail. Reservations are available if you come in and make them.	Took reservation.
H	No reply	Didn't get letter. We have limited space. No reservations are taken. It's hard to answer you. You know Subtown. . . . We don't object but our patrons would. I won't answer on the phone. Come in about it.	Took reservation.
I	Reply sent as result of phone call.	Didn't get letter. [Hostess] We don't like that. Is it absolutely necessary to have them? [Manager]: if it's okay with you I guess it's okay with us.	Took reservation.
J	No reply	Didn't get letter. How many? [10] No reservations for more than 8 on weekends. I will mail you our menu for the following weekend. [Never received.] I'm too busy to look up reservation availability. [Asked on two occasions that we call back.]	Took reservation.
K	No reply	Didn't get letter. I'd think you'd want to come in to discuss something like that. Mixed group? How many people? We can't turn anyone away but never take reservations. You can have a separate room if you come in about it.	Took reservation.

turally established norms varies widely. Discriminatory treatment is minimized when challenged in a direct face-to-face situation,[4] but is maximized when proposals to "violate" group norms are *suggested*. Indirect evidences of discriminatory behavior (subterfuges of various sorts), appear when a direct non-face-to-face challenge is made. At least in the cultural climate of a suburban Northeastern community, the hypothetical Negro is more easily discriminated against than is the Negro-on-the-spot. Certainly we do not expect identical results in different cultural climates (e.g., in the South, or to take a less extreme illustration, the East Side Restaurant Survey of New York City which has shown that Negroes are accorded inferior treatment—poor service, undesirable seating locations, etc.—in many New York restaurants). It may, however, be presumed that the same or related dynamic factors are operating in all three instances.

[4] This hypothesis is supported by a finding of Saenger and Gilbert (3) that a discrepancy between attitudes and behavior occurs where the individual is under high motivation to purchase an article and ignores the fact that the clerk may be a Negro or that the *fait accompli* of a Negro salesclerk makes his presence acceptable. In either case the reality situation appears to act to reduce overt discrimination.

REFERENCES

1. CHEIN, I., DEUTSCH, M., HYMAN, H., & JAHODA, MARIE (Eds.). Consistency and inconsistency in intergroup relations. *J. soc. Issues*, 1949, 5, No. 3, entire number.
2. LAPIERE, R. T. Attitudes vs. actions. *Social Forces*, 1934, 13, 230–237.
3. SAENGER, G., & GILBERT, E. Customer reactions to the integration of Negro sales personnel. *Int. J. Opin. att. Res.*, 1950, 4, 57–76.

Received November 19, 1951.

"REGRESSION" IN SCHIZOPHRENIA AS DETERMINED BY PERFORMANCE ON THE BENDER-GESTALT TEST[1]

BARBARA J. SUTTELL AND GERALD R. PASCAL
American Institute for Research *University of Tennessee*

It is widely contended that one of the primary processes in schizophrenia is regression. Proponents of classical psychoanalytic theory (12, 13, 14) consider schizophrenia a return to the period prior to the structuring of the ego. Several studies, reviewed by Cameron (10), describe parallels in the gross behavior of schizophrenics and of normal children. Gardner (15), for example, attempted to measure regression using a schedule of "psychotic age" in which he equated the behaviors typical of psychotic diagnostic groups with those of normal children of various ages. Several authors have observed that the posturing and the gross motor activity of the catatonic, and the silly activity and incontinence of the hebephrenic, are indicative of a return to childhood. Superficial observation supports this viewpoint. The few studies, however, that attempt using controlled observations report contrary results (11, 18, 19).

In the area of language and concept formation investigators in general agree that in schizophrenia there is a loss, or destruction, of the characteristics of normal adult thought processes accompanied by a reversion to the more concrete forms of thought characteristic of the young child (5, 10). Vigotsky (23) describes schizophrenic regression in this area as the crumbling of an outer layer of abstractive ability (the use of concepts) leaving the primitive nucleus of concrete complexes. Cameron, however, in an extensive series of studies on the language, logic, and concept formation of the schizophrenic as compared with that of the normal child (6, 7, 8, 9), indicates differences as well as similarities in thought processes. He emphasizes a disorganization rather than a destruction of adult thought processes, resulting in a schizophrenic pattern of thought superficially similar yet basically dissimilar to that of the young normal child.

[1] This study was conducted while the authors were working at the Western Psychiatric Institute and Clinic.

Studies of the expressive-motor behavior of schizophrenics invariably indicate regressive tendencies. These studies are numerous and have been reviewed by Anastasi and Foley (1, 2). In the area of perceptual-motor behavior, too, regressive tendencies in schizophrenia have been indicated. Bender states that the most general tendency in the psychotic reproduction of the Bender-Gestalt designs is the reversion to primitive Gestalten (3, 4). She indicates the appearance of loops for dots, preservation of the designs, compression, failure to cross lines, and fragmentation as "regressive" deviations, i.e., the same deviations which occur in the records of young children. Pascal and Suttell (20) also observe that the psychotics are "regressive" in that they tend to reproduce the Bender-Gestalt designs in a manner similar to that of young children.

None of these studies, however, has been designed for a quantitative treatment of schizophrenic regression. In the present study the Bender-Gestalt test performance of schizophrenics is compared with that of young, normal children and with that of psychoneurotics. Comparisons are made in terms of scored items that have been found to discriminate between psychiatric patients and nonpatients, both children and adults (20, 21). The criterion of regressive behavior is defined, for the purposes of the present investigation, as the test performance of young "normal" children, i.e., children between the ages of 6-0 yr. and 9-6 yr., of average intelligence, and not under the care of a psychiatrist.

Procedure

Bender-Gestalt test records were obtained from 54 adult schizophrenics, between the ages of 20 and 45 years, with high school education, and inpatients of a mental hospital. These patients had been diagnosed schizophrenic at staff conference, with no disagreement among participating psychiatrists except in regard to the subtype. All subtypes, except simple schizophrenia, are, however, about equally represented in the group of 54 patients.

The files of the authors contained the following records, obtained from previously reported studies (20, 21).

Adult nonpatient, ages 15 to 50 (high school education)	278
Adult psychoneurotic, ages 15 to 50 (high school education)	117
Children, nonpatient, ages 6 to 9½	297
Children, nonpatient, ages 10 to 14	203

For the two groups of adults 54 records were selected at random from the files. Previous study (20) had shown that, within the age range 15 to 50 years and within average IQ limits, there is no relationship between the variables chronological age, sex, and IQ, and Bender-Gestalt test score. No attempt, therefore, was made to control these variables in selecting the adult records. The 54 records of normal children in each of the two groups were selected at random within the age and IQ categories of the available populations, since previous study (21) had shown a relationship between Bender-Gestalt test score and the variables chronological age and IQ for children. Mean Bender-Gestalt test scores and standard deviations (Table 1) calculated on the basis of the 54 records for each of the three adult groups indicate that the groups are representative of the total populations from which they were selected (20). Comparable statistics are not available for the total populations of the groups of children.

For each of the five groups the frequency of occurrence of each type of scored item was tallied. Items were selected for further analysis on the basis of occurrence of the item in at least four of the five groups of subjects ten per cent or more of frequency of possible occurrence.[2] Frequency of possible occurrence was determined for each item by multiplying by 54 the number of times the particular item could occur in one test record. An item can occur in from one to eight of the designs on each record.

[2] Per cent of possible occurrence was used for inter-item comparisons rather than frequency of actual occurrence because the frequency with which an item can occur differs from item to item. Per cent of possible occurrence is the frequency of actual occurrence of the item in the 54 records of the group, divided by the frequency of possible occurrence for the particular item, multiplied by 100.

Fourteen items met this criterion. Of these 14 items, 12 distinguished between the children six to nine and one-half years of age and the adult nonpatients. These twelve items may be considered indicators of regressive perceptual-motor behavior according to our definition, that is, on each of these items the frequency of deviation is significantly greater in the records of children between six and nine and one-half years of age than in the records of nonpatient adults.

RESULTS AND DISCUSSION

In Table 2 the quantitative data are presented for the 12 items which meet our criterion for regression indicators. As shown in column 1, there is a significant difference in frequencies of occurrence for each of these 12 items between the group of adult nonpatients and the group of younger normal children. Since, according to our definition, these items are indicators of regressive perceptual-motor behavior, it follows that if schizophrenics are regressive, in this sense, the incidence of deviation on these 12 items should discriminate between the schizophrenic and the nonpatient adult groups.

Column 2 of Table 2 shows that only four of these deviations discriminate significantly ($p<.01$) between the schizophrenics and the nonpatient adults; and two others semi-significantly ($.01<p<.05$). A total of six of the 12 regression indicators, then, may be said to discriminate between these two groups.

It may be noted, from columns 4 and 5, that the items that do not discriminate between the schizophrenic and the nonpatient adult groups are those same items (Nos. 5–10) that discriminate between the two groups of children, and not between the older children and the adults. In addition, the reverse is true. The items that discriminate between the schizophrenic and the nonpatient adult groups are those same items that do *not* dis-

TABLE 1

MEAN SCORES AND STANDARD DEVIATIONS FOR TOTAL POPULATIONS AND SELECTED GROUPS

GROUP	MEAN B-G SCORE		*SD*	
	TOTAL	SELECTED	TOTAL	SELECTED
Adult Nonpatient	18.0	20.0	9.4	9.2
Adult Schizophrenic	47.2	45.9	21.3	28.1
Adult Psychoneurotic	34.5	35.4		16.3
Children 6 to 9½ Years		71.9		34.0
Children 10 to 14 Years		43.5		22.1

criminate between the two groups of children, and do discriminate between the older children and the nonpatient adults. Numerous developmental studies (3, 16, 17, 21, 22) indicate that until the age of nine or ten years maturation of capacity is of great importance in perceptual-motor performance. At nine to ten years of age an essentially adult level of performance capacity is reached. On this basis we may assume that those items which discriminate between the two groups of children measure those aspects of performance that depend to a great extent on maturation of capacity. Whereas, those items which do not discriminate between the two groups of children, but between the older children and the nonpatient adults, measure those aspects of performance that depend to a much greater extent on learning. Our findings, then, suggest that where the ability to execute the aspects of the design seems dependent primarily on maturation, the schizophrenic performance is similar to that of the nonpatient adult. Where, on the other hand, the ability to execute the particular aspects of the design seems dependent more on learning, the performance of the schizophrenic is similar to that of the normal child.

The deviations on which both schizophrenics and children differ from nonpatient adults may offer some clue as to the nature of the similarity in performance. This similarity is, on the basis of our findings, something other than capacity. The aspects of performance on which the schizophrenics resemble children—Shape of Circle, Deviation in Slant, Workover, and Tremor—represent relatively minor inaccuracies. These suggest a failure to control, or to regulate, motor movement. Schizophrenic behavior, then, is similar to that of a normal child, not because of regression to an earlier capacity level but because of the disruption of a learned, apparently regulatory, response. The schizophrenic is capable of the execution of the designs, but fails on those aspects of reproduction dependent primarily on this regulation. Similarly, Cameron (9) found, in studying thought processes, that the schizophrenic is capable of abstraction and generalization, but is unsuccessful in using these abilities because of his failure in organization.

We may approach the comparison of the performances of schizophrenics and children in another way. It is possible to rank order the deviations in terms of frequency of occurrence in the two groups, adult schizophrenics and normal children between the ages of six and nine and one-half years. If the schizophrenic is regressive in the sense that he returns to the reproducing of a perceptual stimulus in the manner typical of a young child, there should be a high correlation in terms of frequency of occurrence of these deviations. A rank difference correlation coefficient, rho, of .34 was obtained, indicat-

TABLE 2

Incidence of "Regressive" Deviations Contrasted for Groups

Indicator	Adult Nonpt. vs. Child. 6–9½	Adult Nonpt. vs. Schiz.	Schiz. vs. P.N.	Child. 6–9½ vs. Child. 10–14	Child. 10–14 vs. Adult Nonpt.
1. Shape of circle *	+	+	−	−	+
2. Deviation in slant	+	+	−	−	+
3. Workover	+	+	−	−	+
4. Angles extra and/or missing	+	?	−	−	+
5. Wavy line	+	−	−	+	−
6. Distortion	+	−	−	+	−
7. Angles for curve (des. 6)	+	−	−	+	−
8. Curve not centered (des. 4)	+	−	−	?	−
9. Point of crossing (des. 6)	+	−	−	?	−
10. No order	+	?	?	+	−
11. Asymmetry	+	−	−	−	−
12. Tremor	?	+	−	−	−

* Deviations after Pascal and Suttell (20). Chi-square levels of confidence, significance of differences in incidence: +=1%, −>6%, ?=2–5%.

ing no significant correlation between the incidence of the deviations in the two groups. In over-all performance, then, the schizophrenic is not regressed.

One basic difference between the schizophrenic and the psychoneurotic, according to most of our texts, is that the former is regressive in his behavior. As shown in column 3, the schizophrenics are not significantly differentiated from the psychoneurotics by these regression indicators. Further, rank order of the frequency of occurrence of the indicator deviations in the two groups yields a rank difference correlation coefficient of .95, which is clearly significant. These results suggest that there is no difference between the schizophrenics and psychoneurotics with respect to the incidence of regression indicators, and that on this basis schizophrenics are not to be differentiated from psychoneurotics.

SUMMARY AND CONCLUSIONS

Certain aspects of performance on the Bender-Gestalt test have been contrasted for five groups of 54 *Ss*, adult schizophrenic patients, adult psychoneurotic patients, adult nonpatients, nonpatient children between the ages six and nine and one-half years, and nonpatient children between the ages ten and fourteen years. Twelve deviations from the stimuli were found that discriminated significantly between the adult nonpatient group and the group of younger children. These deviations were considered, for this study, indicators of perceptual-motor regression.

Only four of the 12 regression indicators were found to discriminate at the 1 per cent level of confidence between the adult nonpatient group and the adult schizophrenic group. These items, on which the schizophrenics do "regress," are those that discriminate between the older group of children and the adult nonpatient group, and do not discriminate between the two groups of children. In addition, the items on which the performances of the adult patient and nonpatient groups are similar are, in general, the same items that discriminate between the two groups of children and do not discriminate between the older children and the adult nonpatient groups. The schizophrenic group in our study, then, was "regressive" only on the items the reproduction of which seems dependent on learning to a greater extent than on maturation of capacity.

Our findings do not support the contention that regression is peculiar to schizophrenia. On none of the regression indicator items were the groups of schizophrenics and psychoneurotics in our study differentiated.

It is concluded that (*a*) where schizophrenic perceptual-motor behavior is similar to that of the young child this similarity is based on disruption of a learned motor response, apparently regulatory in nature, and not on a return to an earlier behavioral capacity; and (*b*) regression as defined in this study is not peculiar to the schizophrenic. The findings of this study, in regard to psychomotor performance, corroborate those of Cameron in the area of language and thought, both indicating only a superficial similarity between the performance of adult schizophrenics and of young normal children.

REFERENCES

1. ANASTASI, A., & FOLEY, J. P., JR. A Survey of the literature on artistic behavior in the abnormal: I. Historical and theoretical background. *J. gen. Psychol.*, 1941, **25**, 111–142.
2. ANASTASI, A., & FOLEY, J. P., JR. A Survey of the literature on artistic behavior in the abnormal: IV. Experimental investigations. *J. gen. Psychol.*, 1941, **25**, 187–237.
3. BENDER, L. *A Visual Motor Gestalt Test and its clinical use.* Res. Monogr. 3. New York: Amer. Orthopsychiat. Assoc., 1938.
4. BENDER, L. Principles of Gestalt in copied form in mentally defective and schizophrenic persons. *Arch. Neurol. Psychiat.*, 1932, **27**, 661–686.
5. BOLLES, M. M. The basis of pertinence. *Arch. Psychol. N. Y.*, 1937, No. 212.
6. CAMERON, N. A study of thinking in senile deterioration and schizophrenic disorganization. *Amer. J. Psychol.*, 1938, **51**, 650–665.
7. CAMERON, N. Deterioration and regression in schizophrenic thinking. *J. abnorm. soc. Psychol.*, 1939, **34**, 265–270.
8. CAMERON, N. Reasoning, regression, and communication in schizophrenics. *Psychol. Monogr.*, 1938, **50**, No. 1 (Whole No. 221).
9. CAMERON, N. Schizophrenic thinking in a problem solving situation. *J. ment. Sci.*, 1939, **85**, 1012–1035.
10. CAMERON, N. The functional psychoses. In J. McV. Hunt (Ed.), *Personality and the behavior disorders.* New York: Ronald, 1944.
11. DUBOIS, P., & FORBES, T. Studies of catatonia: III. Bodily postures assumed while sleeping. *Psychiat. Quart.*, 1934, **8**, 546–552.
12. FENICHEL, O. *The psychoanalytic theory of neurosis.* New York: W. W. Norton, 1945.
13. FREUD, S. On narcissism: An introduction. In *Collected Papers*, Vol. 4. Pp. 30–59.
14. FREUD, S. Psychoanalytic notes upon an autobiographical account of a case of paranoia. In *Collected Papers.* Vol. 3.

15. GARDNER, G. The measurement of "psychotic age": Preliminary report. *Amer. J. Psychiat.*, 1931, **87**, 963–975.
16. GESELL, A. *The first five years of life.* New York: Harper, 1940.
17. GOODENOUGH, F. Study in the psychology of children's drawings. *Psychol. Bull.*, 1928, **25**, 272–283.
18. HUNT, J. McV., & COFER, C. Psychological deficit. In J. McV. Hunt (Ed.), *Personality and the behavior disorders.* New York: Ronald, 1944.
19. LANDIS, C. *et al.* Studies of catatonia. *Psychiat. Quart.*, 1934, **8**, 535–552.
20. PASCAL, G. R., & SUTTELL, B. J. *The Bender-Gestalt Test: Quantification and validity for adults.* New York: Grune and Stratton, 1951.
21. SUTTELL, B. J. The development of visual-motor performance in children as estimated by the Bender-Gestalt Test. Unpublished master's thesis, Univer. of Pittsburgh, 1951.
22. TERMAN, L., & MERRILL, M. *Measuring intelligence.* Boston: Houghton Mifflin, 1937.
23. VIGOTSKY, L. Thought in schizophrenia. *Arch. Neurol. Psychiat.*, 1934, **31**, 1063–1077.

Received November 19, 1951.

THE DISCOMFORT-RELIEF QUOTIENT AS A MEASURE OF TENSION AND ADJUSTMENT[1]

ARNOLD MEADOW
University of Buffalo

MILTON GREENBLATT, JULIUS LEVINE, AND HARRY C. SOLOMON
The Boston Psychopathic Hospital and Harvard Medical School

In 1947 Dollard and Mowrer (2) described an ingenious method of analyzing personal documents. The method was proposed as a measure of psychological "tension." It was based essentially on the ratio between the number of units in a written document expressing discomfort and the sum of the units expressing discomfort and comfort or relief. Subsequently, investigators have published studies in which the ratio, designated by Dollard and Mowrer as the "discomfort-relief quotient" (DRQ), has been utilized as a criterion for the evaluation of therapy (1, 4, 10, 12). These workers have uniformly interpreted a decrease in the quotient at the end of therapy both as an indication of a diminution of psychological tension and as an indication of success in therapy.

The increasingly wide use of the DRQ as an index of tension raises the question of whether it is in fact measuring what it purports to measure. The basic assumption of the method is that the conscious verbal expression of discomfort reflects more or less accurately the psychological tension characterizing the subject. Clinical experience suggests that this assumption may be invalid. In some cases the content of the patient's verbalizations appears to reflect the tension noted in behavioral and physiological observations; in other cases there appears to be no relationship between behavioral and physiological estimates of tension and the patient's verbalizations. The present series of studies was designed to test the validity of the DRQ as a measure of tension and adjustment.

Experiment I. Relationship Between the DRQ and Clinical Rating of Tension

Method

The first study concerns the relationship between the DRQ and clinical ratings of tension. The method consisted in deriving a DRQ for each patient from a free verbalization protocol and correlating the scores so obtained with a rating of tension made by a psychiatrist.

Subjects

The study is based on a group of 35 cooperative chronic schizophrenic patients sent to the Boston Psychopathic Hospital for the lobotomy operation.

The ages ranged from 19 to 57 years with an average of 36 years. There were 27 women and eight men. Twenty-four of the patients were diagnosed as dementia praecox paranoid; one was diagnosed dementia praecox catatonic; two were diagnosed dementia praecox simple; six were diagnosed dementia praecox hebephrenic; and two were diagnosed dementia praecox, other types. All patients were studied prior to operation.[2]

Procedure

The DRQ was obtained from the analysis of the protocols of a newly devised, free verbalization test. The protocols for this test are obtained by the following procedure: The patient, facing the examiner, is asked to tell what is on his mind, or what he is thinking about. Whenever there is an interval of silence of one minute, the examiner asks a nondirective question, such as, "What are you thinking about?" or "What is on your mind?" The interview lasts 30 minutes and is transcribed in shorthand.

The transcribed protocol is then divided into "units" as a basis for further scoring, "unit" being defined as an independent clause.

In obtaining the DRQ each unit is classified as showing (*a*) discomfort or dissatisfaction, (*b*) comfort or satisfaction, or (*c*) neutrality. The DRQ is derived by dividing those units indicating discomfort by the total number of units of the protocol.

The test was administered to each patient twice within a two-week period to determine reliability. Rank order correlation of test-retest reliability was +.82. The rank order correlation between two independent scorers of the same protocols was +.77.

The clinical ratings of tension were made by a psychiatrist who spent several hours in interviewing each patient. The ratings were based on an over-all estimate of tension displayed by the patient in the course of these interviews. Reliability was determined by comparing the psychiatrist's ratings of

[1] This study was financed by a research grant from the United States Public Health Service.

[2] Prelobotomy patients were used because the present data are also being utilized as part of a more extensive research project designed to evaluate the effects of the lobotomy operation.

tension with that of a clinical psychologist who also had had several hours of contact with each patient. The rank order correlation coefficient between the two ratings of tension was +.78.

Results

The correlation between the psychiatrist's clinical rating of tension and the DRQ derived from the free verbalization protocol was +.05. The data, therefore, indicate no direct relationship between the verbalization of tension and the over-all clinical tension level of the patient.

In further examining the DRQ's and clinical ratings of tension of individual patients, possible reasons for this lack of correlation were suggested. There was one small group of patients whose hands perspired freely, whose voices quivered tremulously, whose entire mien betrayed a state of intense excitement. These patients received a very high psychiatric tension rating but appeared to be unable to verbalize either discomfort or relief and consequently achieved a relatively low DRQ. Other patients received a similarly low DRQ score but their calm, placid behavior led the clinicians to rate them as very low in tension. One group of paranoid patients received both a high psychiatric tension rating and a high DRQ score. A second paranoid group received an equally high psychiatric tension rating but extremely low DRQ's on the basis of such verbalizations as "There's nothing wrong with the lower parts of my body, they are beautiful; I am the heavenly mother."

It appears clear from these examples that the tension of some patients is directly expressed verbally, in other cases it is not expressed at all, and in still other cases an emotional state is expressed verbally which is the direct opposite of the behavioral picture. When all of these patients are combined into one group the net result is the zero correlation which appears in the data.

Experiment II. The DRQ as an Index of Adjustment or Success in Therapy

The results of the first experiment indicate that the DRQ is not correlated with a clinical rating of tension. The further hypothesis of previous investigators (1, 4, 10, 11), that a relatively low quotient is an indicator of good personality adjustment or therapeutic success, was tested in the second experiment of this investigation.

Method

The method of the second study consisted in correlating the DRQ with a battery of tests designed to measure the degree of personality integration or adjustment of schizophrenic patients. The same group of patients was used for this experiment as in the first study of this report. The DRQ for each patient was also the same one described in the first experiment.

Three types of measures were used to appraise "adjustment." The first was a series of tests of abstract thinking; the second, a newly devised quantitative measure of "looseness of association." The third was a clinical rating of degree of personality integration.

Tests of Abstract Thinking

The Neutral Proverbs Test. The patient is asked to "explain the meaning" of nine familiar sayings, all of which were considered to be relatively neutral in content. Examples: "A new broom sweeps clean."

The Emotional Proverbs Test. This test also consists of nine proverbs. It differs from the neutral proverbs in that the items of the former test all have relatively emotional content. The patient is again asked to "explain the meaning" of each proverb. Example: "The wages of sin is death."

The Neutral Similarities Test consisted of the first eleven items on the Wechsler-Bellevue Scale (Form I) similarities subtest (11).

The Emotional Similarities Test is a parallel set of items devised to cover areas expected to have emotional significance for psychotic patients. Example: The patient is asked, "In what way are murder and suicide alike?"

Object Sorting Test. The patient is asked to give the common characteristics of a group of objects presented to him. There are twelve items, i.e., groups of objects, in the test.

The rationale, content, and scoring methods for all five tests of abstract thinking are more completely described in previous publications (5, 6, 7, 8).

Measure of Looseness of Association

This score is obtained from the free verbalization protocols. In its derivation, the protocols are first divided into independent clauses. The scorers then count the number of times a "break" appeared between two units of thought. A "break" or a "loosening of association" is defined as "any shift in ideas that would not appear in the discourse of a normal person."

Rating of Personality Integration

The clinical rating of personality integration was based on the combined clinical criteria of social withdrawal, divorce of thought content from affect, and deterioration of habit patterns. The ratings were made by a psychiatrist (J. L.) who had observed and interviewed each patient for several hours.

Test-retest and interscorer reliability coefficients

for the abstraction and looseness of association tests have been presented in a previous publication (8). The rank order correlation coefficients for two administrations of the neutral proverbs test was +.94, for the emotional proverbs test, +.84, and for the looseness of association index, +.78. The rank order correlation coefficient between two independent scorers for the neutral proverbs test was +.89, for the emotional proverbs test +.91, for the neutral similarities test +.97, and for the looseness of association index +.90. All correlations were significant at the 1 per cent level.

Results

The correlations between the DRQ and the various measures of "adjustment" are presented in Table 1.

Table 1 indicates positive correlations between all of the measures of abstraction and the DRQ. The correlations are all significant at the 5 per cent level with the exception of the correlation between the object sorting test and the DRQ ($p=.06$).

The DRQ is also correlated positively with the rating of personality integration and negatively with looseness of association. The results thus indicate that the greater the relative amount of discomfort expressed, the less is the impairment in abstraction, looseness of association, and personality disintegration in schizophrenic subjects.

TABLE 1

CORRELATIONS BETWEEN DISCOMFORT-RELIEF QUOTIENT AND MEASURES OF "ADJUSTMENT" OF SCHIZOPHRENIC PATIENTS

MEASURES	NUMBER OF CASES	RANK-ORDER CORRELATION	PROBABILITY
Discomfort-Relief Quotient and Neutral Proverbs	28	.57	.01
Discomfort-Relief Quotient and Emotional Proverbs	27	.41	.05
Discomfort-Relief Quotient and Neutral Similarities	31	.48	.01
Discomfort-Relief Quotient and Emotional Similarities	31	.61	.01
Discomfort-Relief Quotient and Object Sorting	20	.31	.06
Discomfort-Relief Quotient and Rating of Personality Integration	31	.47	.01
Discomfort-Relief Quotient and Looseness of Association Index	30	—.48	.01

DISCUSSION

The results of the first experiment indicated that there was no correlation between the DRQ and tension as judged clinically. In interpreting this datum, a first consideration is that speech is only one of several avenues for expression of tension and may not be the preferred avenue. The clinician's estimate of degree of tension takes into account facial expressions, gross and fine skeletal muscle movements and autonomic reactions—in addition to verbal responses. Every clinician is familiar with the type of schizophrenic patient who, by his constant grimacing, restlessness, and rapid pulse, reveals considerable inner emotional turmoil, though he may not express any of this turmoil in the manifest content of his verbalizations. The breakdown of interpersonal communication which characterizes schizophrenia perhaps makes the spoken word a relatively unreliable indicator of internal state in this group of subjects.

A second possible interpretation would explain the lack of correlation between the DRQ and tension as judged clinically in terms of an anxiety-defence theory. This theory would postulate that the degree to which tension was directly expressed in the verbalizations of a patient is a function of the type and intensity of the defence he utilizes. Thus, on one hand, a relatively mute catatonic patient who receives a high clinical rating of tension may express verbally relatively little discomfort or tension; he consequently obtains a zero DRQ. On the other hand, a "schizoid-affective" type of patient who receives a similarly high clinical rating of tension might express his tension more directly in the content of his verbalizations; he consequently obtains a high DRQ. Within a group of patients with a similar type of defence, there may be, moreover, different degrees of defence, thus creating a greater or lesser verbal expression of tension. A further complication is the possibility of several types of defence existing in the same patient. The many diverse ways in which the verbalizations of a patient may be distorted by defensive processes may account for the absence of a relationship between over-all tension as

rated clinically and tension as expressed in a patient's own verbalizations.

The positive correlation between the relative amount of tension verbalized and the integrity of intellectual and personality organization may also be interpreted in terms of the anxiety and defence theory. Disintegration of intellectual and personality organization may be conceptualized according to this theory as the manifestation of a defence process that prevents the conscious recognition and verbal expression of feelings of discomfort.

Previous workers who have used the DRQ have assumed that it measured tension and that it was related inversely to good adjustment. The present research, which indicates clearly that the assumption does not hold for a clinically schizophrenic group, raises the question of its validity in neurotic patients. The further use of the DRQ in studies of neurotics, without further experimental justification, would appear to be unwarranted.

Summary and Conclusion

Two experiments designed to test the validity of the use of the DRQ as a measure of psychological tension and adjustment are described. Thirty-five chronic schizophrenic patients randomly selected from a prelobotomy schizophrenic group at the Boston Psychopathic Hospital were used as subjects.

In the first experiment, the DRQ's derived from the protocols of a newly devised free verbalization test were correlated with clinical ratings of tension made by a psychiatrist. The test-retest reliability coefficient for the DRQ was +.82; interscorer reliability for the quotient was +.77. The reliability coefficient for two independent raters of tension was +.78. The results indicated a zero correlation between the DRQ and the rating of clinical tension.

In the second experiment the DRQ was correlated with three groups of measures of adjustment—namely, a series of abstraction tests, an index of "looseness of association," and a clinical rating of "personality integration."

Test-retest reliability coefficients for the measures of abstraction ranged from +.89 to +.94; for the looseness of association index the test-retest reliability coefficient was +.78. Interscorer reliability coefficients for the tests of abstraction and looseness of association ranged from +.89 to +.97.

The results indicated a positive relationship between the DRQ and all measures of adjustment.

The results of both experiments are interpreted to indicate that a relatively high DRQ cannot be used as a valid measure of tension but may be used as an indicator of good adjustment in schizophrenia.

To interpret the positive correlation between a relatively high amount of expressed discomfort and relatively good adjustment, we hypothesize that the more serious maladjustments in schizophrenia are accompanied by intensified defensive processes that prevent the conscious recognition and verbal expression of feelings of discomfort.

The assumption of previous investigators that a relatively low DRQ is indicative of low "tension" state, good adjustment, and therapeutic success is challenged by the results of the present experiments.

References

1. Assum, A. L., & Levy, S. J. Analysis of a non-directive case with follow-up interview. *J. abnorm. soc. Psychol.*, 1948, 43, 78–79.
2. Dollard, J., & Mowrer, O. H. A method of measuring tension in written documents. *J. abnorm. soc. Psychol.*, 1947, 42, 3–32.
3. Haigh, G. Defensive behavior in client centered therapy. *J. consult. Psych.*, 1949, 13, 181–189.
4. Kauffman, P. E., & Raimy, V. C. Two methods of assessing therapeutic progress. *J. abnorm. soc. Psychol.*, 1949, 44, 379–385.
5. Meadow, A. Anxiety, concrete thinking and blood pressure changes in schizophrenics. Unpublished Ph.D. thesis, Harvard Univer., 1951.
6. Meadow, A., & Funkenstein, D. Abstract thinking and the status of the autonomic nervous system in schizophrenia. In P. Hoch and J. Zubin (Eds.), *Relation between psychology and psychiatry*. New York: Grune and Stratton, 1951.
7. Meadow, A., Greenblatt, M., & Solomon, H. C. Abstract thinking and looseness of association in schizophrenia. *J. nerv. ment. Dis.*, in press.
8. Meadow, A., Funkenstein, D., Greenblatt, M., & Solomon, H. C. Concrete thinking and blood pressure responses to mecholyl in schizophrenia. *J. nerv. ment. Dis.*, in press.
9. Rogers, C. R. *Client centered therapy*. New York: Houghton Mifflin, 1951.
10. Rogers, N. Measuring psychological tensions in non-directive counseling. *Personal Counselor*, 1948, 3, 237–264.
11. Wechsler, D. *The measurement of adult intelligence*. Baltimore: Williams and Wilkins, 1944.
12. Zimmerman, J. Modifications of the discomfort-relief quotient as a measure of progress in counseling. Unpublished master's thesis, Univer. of Chicago, 1950.

Received November 27, 1951.

INTELLIGENCE FROM A PERCEPTUAL POINT OF VIEW

ARTHUR W. COMBS
Syracuse University

THERE is a growing trend in psychology toward viewing behavior as a function of perception. More and more we have come to understand that the individual's behavior is not so much a function of the physical stimulus as it is a function of his perceptions of the events to which he is exposed. It is the meaning of events to the individual rather than the externally observed nature of events which seems crucial in behavior. As a result, psychologists in increasing numbers are turning their attention to the problems of human perception and are attempting to observe behavior, not from an external point of view, but from the point of view of the individual who is behaving. This paper is an attempt to relate this method of observation to the problem of intelligence. The question we wish to explore in this paper is: "What is the nature of intelligence viewed from a perceptual or phenomenological frame of reference?"

Intelligence as a Problem of Perception

By the term *intelligence* we ordinarily refer to the effectiveness of the individual's behavior. In a personal frame of reference the individual's behavior is described in terms of the perceptions that he can make in his own unique perceptive field. This perceptive field has been called by Snygg and Combs (16) *The Phenomenal Field* and has been defined by them as "the universe of experience open to the individual at the moment of his behavior." In other words, the behavior of the individual will be dependent upon the perceptions that the individual makes in his phenomenal field at the moment of action. The effectiveness of his behavior will necessarily be a function of the adequacy of those perceptions.

If an entity in the perceptive field is vague and ill defined, the behavior of the individual will be correspondingly vague and lacking in precision. Until the child has clearly differentiated that 2 plus 2 equals 4, this function is comparatively meaningless and his behavior in arithmetic is correspondingly inaccurate and ineffective. Thus, the precision and effectiveness of the individual's behavior will be dependent upon the scope and clarity of his personal field of awareness. Intelligence, then, from a perceptual point of view becomes a function of the factors which limit the scope and clarity of an individual's phenomenal field.

The perceptions that could be made of any given situation, such as looking at a stone wall, for example, are, theoretically, practically infinite in number and quality. As a matter of fact, however, we are strictly limited in our perceptions of a stone wall to those which we, as human beings, can make. The perceptions possible to us are only those that people can make. We cannot, for instance, perceive the wall as it would appear to a man from Mars, or from the interior of an atom, or as it would appear to a centipede. What is more, we cannot even perceive it as it would appear to all people. Different people will perceive different aspects of the wall differently, even at the same instant. I can only perceive the wall, and hence behave toward it, in terms of the perceptions that I, as an individual, can make regarding it. I may, for instance, perceive it as a fine, sturdy fence enclosing my property, while a stone mason friend might perceive it as having been poorly designed or as having been built with too little cement in the mortar mixture. The perceptions open to my mason friend are the result of his unique experience. I, not having such experience, am incapable of those perceptions at this moment.

Potential and Functional Perceptions

Before proceeding further with our discussion of the limiting factors in perception, it is necessary for us to pause for a moment to distinguish between potential and functional perceptions. By potential perceptions I mean those perceptions that exist in the individual's unique field of awareness and that, given the right circumstances at any particular moment,

could occur. The fact that a perception is potentially possible to any individual, by no means, however, means that it will occur at the moment of action. Even those perceptions that I can make potentially may not be active for me at any given moment. Potentially, I might be able, for instance, to perceive the wall that we have just been using as an example as a barrier to be gotten over, as an eyesore to be beautified, as composed of 687 bricks costing me $80.27, or as providing pleasant shade on a hot day. These are all potential perceptions I am capable of making about the wall. They will affect my behavior, however, only when they are active or functioning in my field of perceptions. When I am beating a hasty retreat pursued by a neighbor's angry dog, perceptions about the shade, beauty, or cost of the wall, though potential, are not functional in affecting my behavior. I behave only in terms of my functioning perception of the wall as something to get over—and quickly. The fact that particular perceptions may exist potentially in the phenomenal field of an individual is by no means a guarantee that they may exist functionally at the moment of action.

While the potential intelligence of the individual is of interest in judging his capacities, it is practically always a matter impossible to measure with any degree of accuracy. We can only sample those parts of a phenomenal field that *we* happen to feel are important. Obviously the measurement of a person's potential perceptions in these terms is open to extremely grave sampling error and improves in accuracy only as the individuals tested have common experience in the materials chosen for testing. It seems probable that an intelligence test cannot accurately measure the potential differentiations that the individual can make in his phenomenal field. Rather, what we usually measure are the subject's functional perceptions. That is, we measure what differentiations he can make when confronted with the necessity to do so for one reason or another. We may define these functional perceptions as: those perceptions in the field experienced by the individual at the moment of behaving.

From a perceptual viewpoint, if intelligence is the capacity for effective behavior, *the intelligence of an individual will be dependent upon the richness and variety of perceptions possible to him at a given moment.* To understand and effectively to foster intelligent behavior, it will be necessary for us to be concerned with the limiting factors upon the perceptions of an individual. We need to know not only what the individual *could* perceive, but what he *would* perceive at a given moment of behaving.

SOME LIMITING FACTORS UPON PERCEPTION

Physiologic Limitations on Perception

Certainly the physical limitations upon the organism affect the differentiations possible in the phenomenal field. Some forms of prenatal anomalies, like mongolism, microcephalia, and similar disorders, indubitably reduce the level of operation at which the individual can function and seriously impair the ability of the organism to make adequate perceptions. Similarly, there seems good reason to believe that some types of mechanical or disease injury to the central nervous system may result in impaired functioning, such as occurs in cerebral palsy, birth injuries, prefrontal lobotomy, the aftermath of such diseases as encephalitis or, even, in common childhood diseases accompanied by prolonged high fever. Various forms of endocrinopathies, particularly cretinism, also appear to have limiting effects upon differentiational capacity for some individuals. Such physical or biological limitations upon the organism have been widely studied but account for only a small proportion of those persons operating at impaired intelligence levels.

Other less dramatic forms of physical handicaps may also have important effects upon the perceptions possible to the individual, however. This is particularly true of individuals suffering impairment of various sense modalities which may inhibit the clarity or even the existence of some perceptions. We need to remind ourselves, however, that such persons may have as rich and varied a perceptive field within their own limitations as we have within ours. Testing persons living in one frame of reference with tests based on those of another can easily lead us astray, a fact well known to the makers of some tests for the handicapped. The limitations im-

posed upon perception by such physical handicaps as the loss or impairment of locomotion or the use of arms or hands are also important in limiting certain kinds of perceptions. These people experience different, but not necessarily fewer or poorer perceptions of events than so-called "normals."

Perhaps less well recognized in their effects upon perception are such factors as malnutrition, focal infections, and chronic fatigue, which may reduce both the need for and the ability to make adequate perceptions. It is well known in industrial psychology, for example, that fatigued workers are more likely to have accidents, perhaps because of failure to make the right differentiations at the right time. It is conceivable that persons suffering from chronic fatigue over long periods similarly fail to make differentiations useful to them on later occasions.

Certainly such physical factors as these have important effects upon the ability of the individual to make adequate differentiations in his perceptive field. The more dramatic of these have often been recognized and studied. Others, such as the effects of malnutrition, fatigue, and the like, have been less adequately explored. In spite of the lack of research in respect to some of the physical limitations upon intelligence, far more work has been done in this area, however, than in some of those to be discussed below.

Environment and Opportunity as a Limitation upon Perception

The differentiations in the phenomenal field that an individual can make will, of course, be affected by the opportunities for perception to which he has been exposed. To appear in the perceptive field an event must have been, in some manner, experienced by the person who perceives it. Environmental effects upon perception appear to be of two types, actual or concrete and symbolic or vicarious.

Exposure to Actual Environmental Events

In the first place the perceptions possible to any individual will be limited, in part, by the actual environmental factors to which he has been exposed. Eskimos ordinarily do not comprehend bananas, nor African Bushmen snow, since neither has had the opportunity to experience these events in their respective environments. It is not necessary to go so far afield for illustration, however. In our own country our experience with the testing of children in various parts of the nation has shown that perceptions are highly limited by the environmental conditions surrounding the individual. Mountain children, for example, often give bizarre responses on intelligence tests. Sherman and Henry (15) found intelligence test results on such children arranged themselves in order of the opportunities provided by their environment.

There are differences also between the perceptions of rural and urban children, children from the North and children from the South, mountain and valley, seaboard and plains. Nor are such differences confined only to children. Adults, too, are limited in their perceptions by environmental factors. During the war I worked for a time in an induction station receiving men from the mountains of Kentucky, West Virginia, and southern Ohio. An intelligence test in use at this station was composed of a series of five pictures with instructions to the subject to cross out that one of each series of five objects that did not belong with the others. One set of five pictures showed four stringed instruments, a guitar, harp, violin, bass fiddle, and a trumpet. Large numbers of these back country men crossed out the harp because they had never seen one or because "all the others are things in our band." We cannot assume that these men were less able to make differentiations or had perceptive fields less rich than their examiner on the basis of these tests. We can only suggest that their perceptions are different from those who made the test. Presumably, had they made the test and administered it to the psychologist, the psychologist would have appeared rather dull!

Exposure to Symbolic or Vicarious Events

Differentiations may occur in the perceptive field upon a symbolic basis as well as from exposure to an actual event. That is, perceptions may occur in the individual's field through indirect exposure to experience as in reading, conversation, movies, and other means of communication. Although I cannot directly perceive that it is dangerous to expose my-

self to rays from an atomic pile, for example, I can differentiate this notion through what others whom I respect have told me. Ideas and concepts are largely differentiations of this sort and it is probable that many of our perceptions are acquired through a symbolic rather than an actual exposure. Certainly most of our formal schooling falls in this category, which may explain, in part, why so little of it is effective in our behavior.

It will be recognized at once that exposure to events in no sense completely determines the perceptions that the individual will make. Exposure to events is only one of the factors involved in determining whether or not an event will be differentiated. Even with equivalent exposure, the perceptions we make are not alike. Perception is not an all or none proposition but a selective process. The same person in the same situation at different times may perceive quite different aspects of the situation and behave accordingly. The provision of opportunity to perceive is by no means a guarantee that a particular perception will occur, a phenomenon of which teachers are only too aware. The personal field of the individual is always organized and meaningful and, even with exposure to events, only those aspects that have meaning for the individual in his own unique economy will be differentiated with permanence.

The individual in a particular culture perceives those aspects of his environment that, from his point of view, he needs to perceive to maintain and enhance his self in the world in which he lives. This does not mean he makes fewer perceptions than an individual in another culture; he makes only *different* perceptions. Thus, intelligence tests made in one culture and applied in another do not measure the ability to differentiate, nor do they measure the richness of the individual's field. Perhaps what they really measure is no more than the difference between cultures. American-made intelligence tests applied to other cultures generally show the following arrangement of nationality groups in decreasing order; British Isles, Germany, France, Italy, the Balkans, Asiatic countries. It will be noted that these nationality groups are also roughly arranged in order of the degree of commonality with our own culture.

Time as a Limitation of Perception

Differentiation requires time. The richness of perception, therefore, will be in part a function of how long the individual has been in touch with experiences. While it is true that a perception is possible only when confronted by an experience, it is also true that this exposure must be long enough to make differentiation possible. This principle is familiar to anyone who has looked at a painting for a period of time. The perceptions which can be made are almost limitless if one looks long enough.

In thinking of the effect of time upon differentiation, it is necessary for us to keep in mind that we are speaking of the duration of the individual's experience with an event and not of the observer's experience. Thus, while it may appear to an outside observer that an individual is confronted by an experience, from the individual's own point of view, he may have no contact with it whatever. A child may sit in school all day, apparently exposed to the curriculum, but may actually be experiencing and perceiving quite different aspects of the situation. Perception is an internal, individual phenomenon and may be quite different from that of another person, even in the same situation.

Most perceptions that the individual makes are functions of previous differentiations he has made in his phenomenal field. For example, before one can perceive the mechanics of multiplication he must have perceived addition. In the same way, before he can perceive the function of a sand dome on top of the locomotive he must differentiate the fact that locomotive wheels sometimes slip. Clearly this process of differentiation takes time. It seems axiomatic that to make differentiations an individual must have lived long enough to do so, a fact we recognize in the construction of intelligence tests calibrated for various age levels, and which teachers recognize in the concept of readiness.

Differentiations in the phenomenal field seem to be occurring continuously as the organism seeks to satisfy its needs in the myriad situations of life. In this sense, intelligence never ceases to develop but is continuously increasing so long as the individual remains alive and operating. That intelligence seems

to level off at age sixteen or later is probably a mere artifact of our method of observation. So long as the individual remains in school we have at least a modicum of comparable experience which can be tested in different persons. After the school years, when individuals are free to go their separate ways, this modicum of comparable experience rapidly disappears. The older one gets, the more diverse is his experience. Intelligence tests based upon comparability of experience may thus fail to evaluate properly the effectiveness of adults.

The Individual's Goals and Values as a Limiting Factor in Perception

Up to this point in our discussion we have been dealing with factors affecting perception that are widely discussed in the literature and for the most part are well understood. In the remainder of this paper let us turn our attention to several factors less well explored as they appear in a phenomenological setting. The first of these has to do with the effects of the individual's own goals and values as a limiting factor on perception.

From a phenomenological view the individual is forever engaged in a ceaseless attempt to achieve satisfaction of his need through the goals and values he has differentiated as leading to that end. These goals and values may be explicit or implicit, simple or complex, but they are always unique to the personality itself. The goals of an individual will vary in another respect as well. The individual's goals and values may be either positive or negative. That is, in the course of his experience, the person may differentiate some things as matters to be sought, while other things may be differentiated as matters to be avoided. What is more, although there is a considerable degree of stability in the major goals and values of a particular individual, there may be great fluctuations in how some goals are perceived from time to time, depending upon the total organization of the perceptual field at any moment.

The goals and values an individual seeks have a most important effect upon the perceptions he can make. Once goals have been established by the individual they continue to affect his every experience. Thus, the person who has differentiated good music as a goal to be sought, perceives music more frequently. His entire experience with music is likely to be affected. Certainly his experience will differ markedly from the person who has formulated a goal to avoid music at all costs. In the same way the experiences of children who perceive schooling as something to be sought are vastly different from those of children who try to avoid all aspects of schooling. If the fundamental thesis of this paper is accurate, that intelligence is a function of the variety and richness of the perceptive field, then the individual's goals must have a most important effect upon intelligence. A considerable body of research has been accumulating over the past several years, demonstrating this controlling effect of goals and values on the individual's perceptive experience. Such studies as those of J. M. Levine (8), R. Levine (9), Postman (16), and Bruner (3) are fascinating cases in point.

This effect of goals on perception is by no means limited to the subject whose intelligence we wish to measure. It is equally true of the intelligence test constructor. It leads to the very confusing situation wherein the test constructor with one organization of goals perceives certain experiences to be marks of intelligence for another person who may or may not have similar goals. Indeed, the likelihood is that he, almost certainly, does not have similar goals. Intelligence tests thus become highly selected samplings of perception in terms of what the testers consider important. Low scores do not necessarily mean less rich and varied fields of perception; they may mean only fields of perception more widely divergent from those of the examiner. A young man whom the writer tested at an induction center during the war illustrates the point very well. This young man was a newsboy on the streets of a West Virginia city. Although he had failed repeatedly in grammar school and was generally regarded as "not bright," he appeared on a national radio hook-up as "The Human Adding Machine." He was a wizard at figures. He could multiply correctly such figures as 6235941×397 almost as fast as the problem could be written down. He astounded our induction center for half a day with his numerical feats. Yet, on the Binet

Test given by the writer he achieved an IQ of less than 60! People in his home town, who bought his papers, amused themselves by giving him problems to figure constantly. When not so occupied this young man entertained himself by adding up the license numbers of cars that passed his corner. He was a specialist in numbers. Apparently as a result of some early success in this field, he had been led to practice numbers constantly, eventually to the exclusion of all else. This was one area in which a poor colored boy could succeed and he made the most of it. His number perceptions were certainly rich and varied but other things were not. Although he was capable of arithmetic feats not achieved by one in millions, he was classified as dull! I do not mean to argue that variety of perception is unimportan in effective behavior. I do mean to suggest the importance of goals in determining perception.

Cultural Effects on Goals and Perceptions

We have stated here that the richness of the individual's perceptive field is in part a function of the goals he has differentiated as important or threatening to him. But, clearly these goals are themselves the result of the individual's experience. The culture one grows up in deeply affects the goals one holds. Cultures both restrict and encourage, approve and disapprove the formulation of goals in the individual. This selective effect of the culture in large measure determines the goals sought and avoided by the individual. These goals in turn must exert important effects upon the perceptions that become part of the individual's perceptive field.

I remember the Kentucky moonshiner to whom I once administered the Wechsler-Bellevue. This man could not tell me "how many pints in a quart" although he had certainly been taught this fact in his early schooling. Knowing that my client did a considerable business in bootleg liquor, I framed the question differently and asked "Well, how do you sell your liquor?" He smiled tolerantly and replied, "Oh Boss, I just sell it by the jug full!" In his community to have done otherwise would have been to risk bankruptcy! In a culture where a jug is standard container for spirits, what need to know about quarts?

It is conceivable that low intelligence may be, at least in part, no more than a function of the goals an individual is striving to reach in achieving his need satisfaction. The well-known phenomenon in which intelligence tests give best results in the school years, when experience and goals have a degree of commonality, and break down badly following those years would seem to corroborate this point. Perhaps by concerning ourselves with human goals we can affect perception, and thus intelligence, much more than we believed possible. Can it be that the child of low apparent intelligence is not so much a problem of an unfortunate heredity as an unfortunate constellation of goals or values? We could do a great deal about intelligence if that were true.

The Self-Concept as a Factor Limiting Perception

We are just beginning to understand the tremendous effects of the individual's concept of self upon his perceptions and behavior. Lecky (7), for instance, reports the effect of a change in self-concept in improving the ability of children to spell. Other researches have reported similar effects of the self-concept upon the perceptions which the individual may make. Clinical experience would tend to bear out such observations. Any clinician is familiar with numerous instances in which a child's conception of his abilities severely limited his achievement, even though his real abilities may have been superior to his perception of them. One needs but to go shopping with one's spouse to discover again how one's conception of himself as a male or female affects the things he sees and the things he hears.

Perception is a selective process and the conception one holds of himself is a vital factor in determining the richness and the variety of perception selected. It makes a great deal of difference, for example, how one perceives the president of our country if one conceives of himself as a democrat, a republican, or a communist. One needs but to observe a group of children to become aware that little boys perceive things quite differently from little girls. Professors do not perceive like truck drivers, although when I have had to ride with professor automobile-drivers, I have often wished they did. Thousands of

people in our society avoid perceptions having to do with mathematical functions by their firm concept of themselves as people who "cannot do mathematics." The self-concepts we hold have a very vital effect in selecting the perceptions which become part of our perceptive fields. If the effectiveness of behavior is dependent on our perceptive fields, it follows that the self-concepts we hold must affect the "intelligence" of our behavior.

There is another factor in the effect of the self-concept upon perception that makes it even more important as a selector of experience. That factor is the circular effect of a given concept of self. Let us take, as an example, the child who has developed a concept of himself as "unable to read." Such a child is likely to avoid reading and thus the very experience which might change his concept of self is by-passed. Worse still, the child who believes himself unable to read, confronted with the necessity for reading, is more likely than not to do badly. The external evaluation of his teachers and fellow pupils, as well as his own observations of his performance, all provide proof to the child of how right he was in the first place! The possession of a particular concept of self tends to produce behavior that corroborates the self-concept with which the behavior originated.

Every clinician has had experience with children of ability who conceive of themselves as unable, unliked, unwanted, or inacceptable and perceive and behave in accordance with their perceptions. And this effect is not limited to children alone. It seems to me one of the great tragedies of our society that millions of people in our society perceiving themselves as able to produce only *X* amount, behave in these terms. Society, in turn, evaluates them in terms of this behavior and so lends proof to what is already conceived by the individual. Compared to this waste of human potential in our society, our losses in automobile accidents seem like a mere drop in the bucket. It is even conceivable in these terms that we create losses in intelligence. If, in our schools, we teach a child that he is unable and if he believes us and behaves in these terms, we need not be surprised when we test his intelligence to discover that he produces at the level at which we taught him!

It is conceivable that psychology has unwittingly contributed to this situation by the widespread publication of a static conception of intelligence and human capacities. The concept of severe limits upon the capacities of the organism simply corroborates the self-concept of the man in the street and decreases the likelihood of change in his concept of self. Even more important must be the effect upon our educational system. Teachers who believe in an unchanging character of child capacities provide the attitudes and experiences that produce and maintain a child's conception of self and his abilities. It is notorious that children's grades vary very little from year to year through the course of schooling. This continuous and little-changing evaluation must have important effects on the self-concept of the child. If the school system in which the child lives is thoroughly imbued with the notion that a child's capacities are comparatively fixed, it is even conceivable that the system may in large measure produce a child's intelligence level by the circular effect we have mentioned above.

Threat as Factor in Perception

The last of the factors I should like to discuss as a possible factor in intelligence is the effect of threat upon the perceptive field. If our fundamental assumption that intelligence is a function of the richness and breadth of the phenomenal field is correct, the effect of threat on this field becomes a most important consideration. Although these effects have been so widely understood by the layman that they have been made a part of his every day speech, it is interesting that until very recently the phenomenon has been given little attention by psychologists. The perception by the individual of threat to himself seems to have at least two major effects upon the perceptive field.

Restriction of the Perceptive Field Under Threat

The first of these effects is the restrictive effect that the perception of threat to self seems to have on the individual's perception. When he feels himself threatened, there appears to be a narrowing of the perceptive field to the object of threat. This has often been described in the psychology of vision as "tunnel vision." The phenomenon is ex-

tremely common and almost everyone has experienced it at some moment of crisis in his lifetime. One hears it described in such comments as "All I could see was the truck coming at us," or, "I was so scared I couldn't think of a thing." There seems reason to believe that this effect is not limited to traumatic experiences alone, but exists in lesser degree in response to milder threats as well. Combs and Taylor (4), for example, have demonstrated the effect under extremely mild forms of threat.

Such limiting effects on perception must certainly have a bearing upon perceptions available to the individual in his phenomenal field. Subjects who have participated in food deprivation experiments report uniformly that when threatened by hunger, food becomes an obsession (13, 14). Recently, at the dinner table, I asked my young daughter what she had learned at school that day. "Oh nothing," said she with much feeling, "But was our teacher mad! Wow!" It would appear from her remarks that, feeling threatened by an angry teacher, it was difficult for her to perceive much else. Her perceptions of the day were apparently entirely concerned with the nature of anger. No doubt these are valuable perceptions to possess, but I know of no intelligence test which measures them.

I recall, too, the behavior of two little girls whose mother was taken to a mental hospital at the beginning of the summer. The matter was kept a deep secret from these two children for fear they "would not understand." The children spent most of the summer with the writer's daughter in an incessant game of "hospital." From morning to night this game went on outside our living room window. Apparently, this preoccupation was the direct outcome of the threat they felt in the loss of their mother, for with the mother's return the game ceased as suddenly as it had begun. To the best of my knowledge it has not occurred since. Under threat there seem to be severe limits imposed upon the breadth and character of perception.

Defense of the Perceptive Field under Threat

There is a second effect of threat upon the individual's perceptions. This effect has to do with the defense reactions induced in the individual on perceiving himself to be threatened. The perception of threat not only narrows the field and reduces the possibility of wide perceptions, but causes the individual to protect and cling to the perceptions he already holds. Thus, the possibility of perceptual changes is reduced and the opportunities for new perceptions or learning are decreased. Under threat, behavior becomes rigid. The fluidity and adaptation which we generally associate with intelligent behavior is vastly decreased. A number of interesting experiments in the past few years have demonstrated this phenomenon. Cowen (5), for example, illustrated this effect in problem solving.

Our own experiment previously mentioned also demonstrated this effect with even very mild forms of threat. This rigidity or resistance of perception to change under threat is well known to the layman and is well illustrated in some of the sayings of our culture. Such aphorisms as "Nobody ever wins an argument" or "You can lead a horse to water but you cannot make him drink" seem to be illustrations of a vague understanding of the phenomenon in the public mind. It is surprising that this principle has been so long overlooked.

I think it will be generally agreed that intelligent behavior is quite the antithesis of rigidity. In the terms we have used in this article, intelligent behavior is a function of the variety and richness of perception in the phenomenal field. Whatever produces narrowness and rigidity of perception becomes an important factor in limiting intelligence. If this reasoning is accurate, or even partly so, one is led to wonder about the effects of long continued threat upon the development of intelligence. What of the child who has suffered serious threats to himself for long periods of his life, as in the case of the delinquent, for example? Or what of the child who has been seriously deprived of affection and warmth from those who surround him over a period of years? Is it possible that we have created low intelligence in such children? Axline (1, 2) has reported a number of cases in which intelligence scores improved considerably under therapy. We have observed similar changes in our own clinical practice.

It may be argued that, although threat seems to reduce perception, some people under threat apparently produce more effectively. I think, however, it is necessary for us to distinguish between "threat" and "challenge." In threat, the individual perceives himself in jeopardy and feels, in addition, a degree of inadequacy to deal effectively with the threat perceived. In challenge, the individual perceives himself threatened but feels at the same time a degree of adequacy to deal with the threat. It would appear that whether an event is perceived as threatening or challenging is a function of the individual's feeling of competence to deal with it. If this analysis is correct, it would explain why a situation that appears threatening to a person, from the viewpoint of an outside observer, might one time produce rigidity and another highly effective behavior. This description of events seems characteristic of the history of civilization as well as of individuals, if Toynbee's (17) explanation can be given credence. He points out that the most productive (more intelligent?) societies are those in which the society faces some crisis within its capacities to cope with the situation (challenge), while societies without crisis or in which the crisis is overwhelming produce very little or collapse entirely.

Some Implications of this Conception of Intelligent Behavior

If the conception of intelligence we have been discussing in this paper should prove accurate, it seems to me to raise serious questions about some of our common assumptions with respect to intelligence and, at the same time, opens some exciting new possibilities for the treatment or education of persons we have often assumed to be beyond help. It implies that our conception of the limiting factors of intelligence may have been too narrow. It would suggest perhaps that our very point of view with respect to intelligence may have resulted in our own tunnel vision, such that we have not been able to perceive other factors given little attention to this point. Perhaps we have been too impressed with the limitations upon growth and development which we observe in physical maturation. We may, for instance, have jumped too quickly to the assumption that intelligent behavior was limited as severely as physical growth and that we have explored to exhaustion other factors that may limit intelligence.

I am not suggesting that physiologic limits do not exist in respect to intelligence. I am suggesting that we may have conceded too early that we had approached those limits. There is no doubt that we can demonstrate in some cases, such as mongolism, cretinism, and the like, that physical factors severely limit intelligence. But these cases are comparatively few compared to the so called "familial" cases of low intelligence that we often assume are hereditary in origin. What evidence do we really possess that would lead us to the position that an individual of "normal" physical condition and vigor may be limited in his capacity for effective behavior by some physical condition? We assume there must be such factors operating because we cannot explain his handicap otherwise. That biological science has not yet been able to demonstrate such physical bases has not deterred us in this. On the contrary, we have simply deplored the lack of sufficient advance in that discipline to demonstrate our conclusion! I should like to suggest that this may not be their failure but ours. Until it can be definitely established that limitations exist as biological functions, our task as psychologists is to assume that they may just as well be social or psychological in character and to work just as hard exploring the matter in our discipline as we expect the biologist to work in his.

Let us, for example, explore to the very fullest the possibility that in those cases where we cannot demonstrate biologic impairment, the limitations upon intelligence may be psychological. If it turns out not to be true, we shall find out in time. I do not believe we can afford to limit the places where we look by the pre-perceptions we have about the matter. Our responsibility here is too great. Education, to name but the most obvious of our social institutions, has in large measure predicated its goals and methods on a concept of humanity with certain static limitations on intelligence. If these limitations are not static, it is up to us as psychologists to find out. The task of the scientist is to question, not to be content with answers. We cannot afford to accept an undemon-

strated point of view that prevents us from asking questions.

Some Implications for Intelligence Testing

If the concepts of intelligence we have been discussing prove accurate, another area of psychological thought toward which we must cast a quizzical eye is the area of intelligence testing. This is particularly important at a time when our culture has come to accept these instruments as trustingly as the family doctor's prescription. If our approach to intelligent behavior as a function of the variety and richness of the perceptual field is a valid consideration, we need to ask regarding these tests at least the following questions:

1. Is our sampling of the perceptive field truly adequate? If I lived for years in a prison cell, I presume I should become expert in perceptions about that cell. Unfortunately, they would be of little value outside the prison walls, but can it truthfully be said that my perceptions are less rich or varied, or only that they are less rich and varied about things I have not had opportunity to experience? Is the delinquent, with rich and varied perceptions on how to elude the police, less intelligent or has he simply not perceived things society wishes he had?

2. Since perceptions are always closely affected by need, by whose need shall we sample perceptions—yours, mine, society's, the subject's own? I suspect that in terms of his own needs and perceptions the subject might be deemed quite brilliant, though he might or might not appear so from the point of view of society. For the most part our tests are based on the assumption that academic, upper middle-class, intellectual perceptions are important. But are they? Can we assume that the expert machinist, who can perceive things "out of this world" for most of the rest of us about a piece of stock on his lathe, is less intelligent than a diplomat who perceives many things about foreign affairs? Can we be so sure of our values as to call one bright and the other dull? Can we blame the machinist for his lack of perception about foreign affairs without asking the diplomat to be equally skilled in the machinist's field of perceptions?

3. Finally, if perceptions are affected by the factors we have discussed in this paper, is it fair to sample intelligence irrespective of the control of such factors? Shall we, for example, examine the child who has lacked opportunity to perceive, has possessed a concept of self or been so threatened over a long period of time so as to have been unable to perceive what we wish to sample without consideration of those factors? Shall we overlook such factors and be satisfied that the perceptions important to us are not there, or shall we seek for ways to make it possible for the child to have them? Shall we assume that our failure to discover a particular perception present in the field is, *ipso facto*, evidence of lack of capacity; or seek to discover why it is not? On the positive side of the picture, if the concepts we have here been discussing are sound, there is reason to believe that intelligence may be far less immutable than we have thought. It may be that we can do far more than we have dreamed we could. Perhaps we may even be able to create intelligence!

Implications for Constructive Action

Who can say, for example, what results we might be able to achieve by a systematic effort to remove or decrease the effectiveness of the limitations on perception discussed in this paper? It is fascinating to speculate on the possibilities one might try in constructing a situation for a child, or adult, consciously designed to minimize the limitations imposed on perception by physical condition, environment, goals, the individual's self concept, and the effects of perceived personal threat.

If the position we have taken is accurate, it would suggest that there is much we can do (*a*) to free individuals from the restraints upon perception and (*b*) to provide the opportunities for perception to occur.

1. First and most obviously, we should be able to discover and make available to far more people the means to achieve better physical condition. We have already done a good deal in this area but much needs yet to be done. Who can say, for instance, what completely adequate medical care for all our people might mean a generation hence?

2. If this discussion has merit, there lies the possibility of providing experiences for people that will make adequate perceptions possible. We have tried to do this in our schools, but

have not always accomplished it. We have succeeded very well in gathering information and in making it available to students. We have not succeeded too well in making such information meaningful. Can it be that the decreases in school success with advance through the school years is more a function of lack of meaning for students than lack of intelligence? Is it enough to assume that experience provided by us to the student is truly provided when he is free to experience it? Has the child in school, who is so worried about his relationship with his peers that he cannot perceive what his book is saying, truly been provided opportunity to perceive?

In our training of children of "low intelligence," we often provide situations wherein they are carefully taught to perform repeatedly a simple act. Is it possible that in so doing we may be further narrowing their fields of perception and building self concepts that produce even narrower perceptive fields?

What kinds of environments could we construct that might more effectively result in increased perception? Such experiments as Lippitt and White (10) have carried on with democratic and autocratic environments suggest some possibilities, but we need to know much more. Perhaps we could learn to build such environments from observing with greater care and understanding the methods of good teachers.

3. Who can say what possible effects might occur from a systematic release of the individual's perceptions by the satisfaction of his most pressing needs or goals? We college professors insist we can produce more, which is another way of saying perceive more, when we have the leisure time to do so, when we are freed from the necessity of spending our time satisfying our needs for sheer existence. Can this be less true of others? It is possible that the child with needs of love, affection, status, prestige, or a girl friend might also be freed to perceive more widely and richly, if we could but find ways of helping him satisfy his needs. Ordinarily, we pay a good deal of attention to the physical needs of a child, understanding that with these needs unfulfilled, he makes a poor student. Is there any good reason to suppose his psychological needs are less pressing or less important in freeing him to perceive widely and accurately? We spend much time and energy trying to find ways of "motivating" people or blaming them for not being motivated to do what we need them to do. We assume that if permitted to seek their own needs, people will not satisfy ours. Perhaps we should get further by helping them satisfy their needs; they might then be free to satisfy ours.

4. Most of our educational methods are directed at the provision of perceptions for the student. He is lectured, required, shown, exhorted, and coerced to perceive what someone thinks he should. It seems possible that with equal energy devoted to the matter of creating needs, goals, and values in students, rich and varied perceptions might be more efficiently produced.

What effects might we be able to produce by providing experiences that build adequate concepts of self in children and adults? What differences in the richness and variety of perception might result from a generation of people with "I can" rather than "I can't" conceptions of themselves? What possibilities of increased perceptions and hence of increased intelligence might accrue to such a program? Clinical experience has demonstrated frequently how a changed perception of self as a more adequate personality can free children for improved school performance, for example.

What would happen if we were consciously and carefully to set about the task of providing experiences that would lead people to conceptions of themselves as adequate, worthy, self-respecting people? If freedom to perceive is a function of adequate perceptions of self, it should not surprise us that the child who perceives himself as unwanted, inacceptable, unable, or unliked behaves in rigid fashion. It should be possible, too, to reverse this process and produce more adequate perceptions by systematic efforts at producing more adequate definitions of self. The possibilities seem tremendous but we have scarcely scratched the surface of this problem.

Finally, if threat to the individual has as important effects as seem indicated in this discussion, the removal of threat would seem a most important factor to consider in the release of the individual to perceive more

adequately. The work of Rogers (12) and his students in client centered therapy has already illustrated to some degree what possibilities freeing the individual to perceive more adequately may accomplish through the provision of a permissive nonthreatening relationship between counselor and client. We have already mentioned the effects Axline (1, 2) has reported following a permissive, nonthreatening form of play therapy.

Such effects do not seem limited to the therapeutic situation, however. A number of workers have applied this principle of permissiveness to the classroom situation with equally gratifying results. Experiments in student centered teaching at Syracuse (6) have led many of us to believe in the tremendous educational possibilities in the removal of threat.

This paper has asked many questions. Indeed, it has asked far more questions than it has presumed to answer. That, it seems to me, is the function of theory. The picture of intelligence presented here as it seems from a phenomenological viewpoint may be accurate or false or, more likely, partly true and partly false. Only time and the industry of many observers can check its adequacy or inadequacy. It seems to me to pose problems that are both exciting and challenging. If it proves as stimulating to the reader as it has to the author, I shall rest content that a theory has achieved its purpose.

REFERENCES

1. AXLINE, VIRGINIA M. Mental deficiency—Symptom or disease? *J. consult. Psychol.*, 1947, 13, 313–327.
2. AXLINE, VIRGINIA M. Play therapy: A way of understanding and helping "reading problems." *Childhood Educ.*, Dec. 1949.
3. BRUNER, J. S., & GOODMAN, C. G. Value and need as organizing factors in perception. *J. abnorm. soc. Psychol.*, 1943, 38, 507–517.
4. COMBS, A. W., & TAYLOR, C. The effect of perception of mild degrees of threat on performance. *J. abnorm. soc. Psychol.*, 1952, 47, 420–424.
5. COWEN, E. L. A study of the influence of varying degrees of psychological stress on problem solving rigidity. Unpublished doctoral dissertation, Syracuse Univer., 1950.
6. LANDSMAN, T. An experimental study of a student-centered learning method. Unpublished doctoral dissertation, Syracuse Univer., 1950.
7. LECKY, P. *Self consistency.* New York: Island Press, 1945.
8. LEVINE, J. M., & MURPHY, G. The learning and forgetting of controversial material. *J. abnorm. soc. Psychol.*, 1943, 38, 507–517.
9. LEVINE, R., CHEIN, I., & MURPHY, G. The relation of the intensity of a need to the amount of perceptual distortion: A preliminary report. *J. Psychol.*, 1942, 13, 283–293.
10. LIPPITT, R., & WHITE, R. K. The "social climate" of children's groups. In R. G. Baker, J. S. Kounin, and H. F. Wright (Eds.), *Child behavior and development.* New York: McGraw-Hill, 1943.
11. POSTMAN, L., & SCHNEIDER, B. H. Personal values, visual recognition and recall. *Psychol. Rev.*, 1951, 58, 271–284.
12. ROGERS, C. R. *Client centered therapy.* Boston: Houghton Mifflin, 1951.
13. SANFORD, R. N. The effects of abstinence from food upon imaginal processes: A preliminary experiment. *J. Psychol.*, 1936, 2, 129–136.
14. SANFORD, R. N. The effects of abstinence from food upon imaginal processes: A further experiment. *J. Psychol.*, 1937, 3, 145–159.
15. SHERMAN, M., & HENRY, T. R. *The hollow folk.* New York: Crowell, 1933.
16. SNYGG, D., & COMBS, A. W. *Individual behavior.* New York: Harper, 1949.
17. TOYNBEE, A. *A study of history.* New York: Oxford Univer. Press, 1947.

Received November 27, 1951.

A COMPARISON OF THREE EXTINCTION PROCEDURES FOLLOWING HEART RATE CONDITIONING

J. M. NOTTERMAN,[1] W. N. SCHOENFELD, AND P. J. BERSH[2]

Columbia University

THE present report concerns an experimental comparison of three different extinction procedures used with a conditioned response of the human heart. The heart conditioned response (CR), a depression in rate, was established following a method described earlier (4). This method is based on operations characteristic of laboratory procedures used for the development of "experimental anxiety" (6). One focus of interest in the case of such conditioned autonomic behavior lies in the changes in strength of these CR's under various post-acquisition treatments. Accordingly, three extinction procedures were tried: (*a*) simple extinction in which the reinforcing stimulus (in this instance, electric shock) was omitted; (*b*) prior verbal instruction to the subject that the reinforcing stimulus is to be omitted, plus the actual omission of it; and (*c*), providing the subject with an avoidance response which would, according to verbal instruction, forestall the reinforcing stimulus.

METHOD

Essentially, the apparatus, described elsewhere (4), consisted of a cardiograph, shock circuit, and tone-oscillator circuit. The shock, 30 volts a.c., was administered by means of two electrodes strapped to the palmar and dorsal sides of the left hand. The subject (*S*) received a tone of 750 cps, approximately 20 db, through a headset. The tone was presented for a duration of one second, and the shock for six seconds. On trials during which shock followed tone, a six-second interval separated the paired stimuli. There were three phases of the experiment, occurring in the following order: (*a*) determination of basal (preconditioning) heart-rate response to the tone, (*b*) conditioning, and (*c*) extinction. The first two phases were the same for all of the 30 male college students who served as *S*s in the experiment. For extinction, however, these *S*s were randomly assigned to one of the three different procedures.[3] All phases were administered in a single session lasting approximately 105 minutes.

As *S* entered the laboratory, he was asked to strip to the waist, and to sit in a chair. The experimenter (*E*) then placed the cardiograph lead over the apex of *S*'s heart (lead IV of EKG), with the "free" electrode strapped to his left foot. A third electrode was used as a ground, and was placed over the xiphoid process. While this was being done, the following remarks were made to *S*:

> This is an experiment to determine your heart-rate reaction to tone, and your heart-rate reaction to electric shock. You will hear a series of tones through this headset which will soon be placed over your ears. A little later in the experiment, you will receive a series of electric shocks administered to your left hand; you will not be shocked through the electrodes I am now placing on your chest and left leg. These electrodes enable me to record your heart action. The shocks you will receive will be strong enough for you to feel them, but will not be strong enough to do you any harm. Remember that they will just pass through your hand, and not your body. Please try not to make any unnecessary movements. I will be keeping a continuous record of your heart action while you are in here, and since the apparatus is necessarily very sensitive, your moving around too much would interfere with the recording.

The *E* then left *S*'s room and entered the apparatus room, some 15 feet away. *S*'s room was sound-resistant; inadvertent apparatus cues were minimal. A pair of leads in a shielded cable were run from *S* to the voltage amplifier of the cardiograph. The basal trials were begun immediately thereafter; a series of 20 one-second tones, spaced irregularly one to two minutes apart, was sent through the headset, *S*'s cardiogram being taken continuously from about 30 seconds before to about 30 seconds after each tone. Heart rate measures were taken by determining the period of the last two heart cycles immediately preceding onset of tone, and the period of the last two cycles in the six-second interval following tone (and, therefore, immediately preceding shock on the subsequent tone-shock trials). These measures, called the "pretone" and "posttone" heart rate, respectively, were then converted to beats per minute.

Following the basal series, conditioning was begun immediately for all *S*s without further instruction. This phase consisted of 18 trials, of which 11 were tone-shock pairings, and seven were tone-alone trials. The reinforcement schedule of these 18 trials was as follows: the first five in the series were tone-shock, the remaining 13 were randomly interspersed tone-shock trials (six) and tone-alone trials (seven). The randomization of reinforced and unreinforced trials was not begun with the first trial, because it has been found that responses become conditioned more readily under partial reinforcement schedules if some response strength is built up initially with regular reinforcement (3).

[1] Postdoctoral Research Fellow, National Institute of Mental Health, U. S. Public Health Service, 1950–52.

[2] Postdoctoral Research Fellow, National Research Council, 1950–51.

[3] Some of the data relating to the performance of ten of these *S*s (the non-instructed group, to be described shortly) have been reported in another paper (5).

It may be noted here that the first occasion on which tone is followed by shock (trial 1 of conditioning) is, as far as *S*'s heart-rate response to tone is concerned, just another basal trial. It is only upon the second trial of conditioning that the tone can be considered as a stimulus in the process of developing CS properties. For this reason, in the computations and graphs reported, the first trial of conditioning has been treated as still another basal trial. Similarly, the first trial of extinction has been treated as the last trial of conditioning. In addition, this makes possible a cross comparison with data previously reported (4, 5), which had been similarly treated.

The three extinction groups to which *S*s were then randomly assigned were treated as follows: Group I, the "non-instructed" group, was started on extinction without being so informed. Eleven tone-alone trials, spaced one to two minutes apart, were administered to this group. Group II, the "instructed" group, was read the following statement, immediately following the first extinction trial: "From now on, you will not be shocked any more following tone presentation." A series of 10 additional tone-alone trials, as for Group I, was then presented. Group III, the "instructed-avoidance" group, was read the following statement immediately following the first extinction trial: "From now on, you will not be shocked any more following tone presentation, provided you tap this key as soon as you hear the tone. Keep your hand near the key." A telegraph key was then put in position on the table, near *S*'s right hand. A simple adjustment, requiring but a second or two, placed the key in a circuit; pressure on the key closed an electric contact, and resulted in a mark being made on the EKG tape. Ten more tone-alone trials were then given, as for the previous groups.

RESULTS

A trial-by-trial plot of the acquisition and extinction of the conditioned response is given in Figures 1 and 2, respectively. The former is based on the trial-by-trial means of the pretone minus posttone heart rate differences for all 30 experimental *S*s; the latter gives these means for the 10 *S*s in each of the three extinction groups separately.

The data entering into Figures 1 and 2 are shown in Figure 3 in more compact form. The points to the left of the broken vertical line represent acquisition data pooled by sections of three trials. Thus, point B–1 is the mean of trials 19 and 20 of the basal series, and trial 1 of conditioning; this point is therefore a mean of 90 readings—three trials for each of 30 *S*s. Point 17–E is the mean of trials 17 and 18 of conditioning, and trial 1 of extinction. The data to the right of the broken vertical line are for extinction; the last point is for trials 8, 9, 10, and 11.

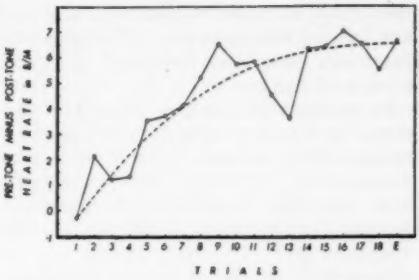

FIG. 1. ACQUISITION OF CONDITIONED HEART-RATE RESPONSE UNDER A PARTIAL REINFORCEMENT SCHEDULE. POINT E IS THE FIRST TRIAL OF EXTINCTION.

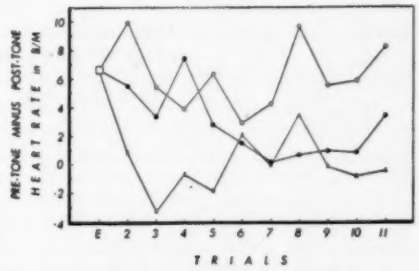

FIG. 2. EXTINCTION OF CONDITIONED HEART-RATE RESPONSE UNDER THREE DIFFERENT PROCEDURES. OPEN CIRCLES, NON-INSTRUCTED GROUP; SOLID CIRCLES, INSTRUCTED GROUP; TRIANGLES, INSTRUCTED-AVOIDANCE GROUP.

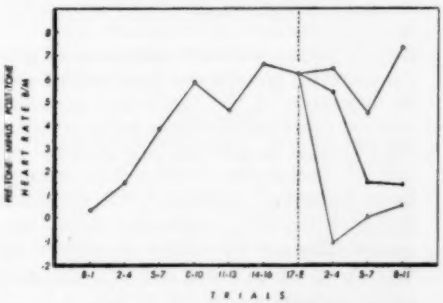

FIG. 3. ACQUISITION AND EXTINCTION OF CONDITIONED HEART-RATE RESPONSE. EXTINCTION SYMBOLS ARE AS FOR FIG. 2.

The data of Table 1 afford a statistical comparison of the different extinction procedures. For each *S* the mean difference in pretone minus posttone heart rate was determined for the first five and the last five trials of extinction. Intergroup comparisons were

then made in terms of these first five trial and last five trial measures. Wilcoxon's nonparametric tests for significance (7) were used in all cases.

In general, the curve of acquisition presented in Figure 1 is fairly typical of conditioning curves normally secured with other responses (2). A comparison of conditioned heart rate data resulting from regular vs. irregular reinforcement techniques is made elsewhere (5).[4] Of chief interest in the present study are the extinction curves shown in Figure 2. The characteristics of these functions may be summarized as follows: the non-instructed group reveals little tendency, if any, to decrease in response strength; the instructed group shows a gradual decrease in magnitude of the conditioned response; and the instructed-avoidance group reveals a marked drop in level immediately after extinction commences, with a low strength being maintained thereafter.[5] The extinction phase of Figure 3 emphasizes these intergroup differences by pooling the data of the first three, second three, and last four trials of each group.

Intergroup comparisons afforded by the data of Table 1 reveal the following chief features of interest:

Non-instructed vs. instructed. The mean pretone minus mean posttone values for the two groups are not significantly different for the first five trials of extinction, but are different at the .01 level for the last five. This finding receives support from previous studies involving the variable of verbal information prior to extinction (e.g., 1).

Non-instructed vs. instructed-avoidance. This comparison indicates a difference significant at the .01 level for the first five trials, as well as for the last five trials. The addition of an avoidance response to the verbal information factor apparently results in a much more rapid extinction than is the case with instructions alone.

Instructed vs. instructed-avoidance. The direct comparison made here serves further to emphasize the effect of the avoidance response. The difference between the means for the first five trials is significant between the .05 and .02 levels. The data for the last five trials are not significant. Thus, the chief effect of the avoidance response apparently takes place during the initial stages of extinction.

TABLE 1

INDIVIDUAL AND GROUP MEAN PRETONE AND MEAN POSTTONE HEART RATES (IN BEATS PER MINUTE) FOR THE FIRST FIVE AND LAST FIVE EXTINCTION TRIALS

SUBJECTS	FIRST FIVE TRIALS *						LAST FIVE TRIALS *					
	GROUP I		GROUP II		GROUP III		GROUP I		GROUP II		GROUP III	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
1	75.7	63.8	72.2	60.2	78.7	80.9	81.2	62.5	69.1	67.3	77.5	77.3
2	87.9	85.8	74.3	75.1	62.8	68.6	89.6	82.0	71.3	70.9	63.6	64.1
3	71.5	69.3	85.5	84.5	69.0	70.1	69.6	66.9	88.7	87.0	71.3	70.9
4	87.4	83.7	77.8	73.6	59.8	62.5	90.5	85.4	72.7	69.9	62.3	62.0
5	84.1	74.3	77.8	79.4	76.3	70.2	81.8	75.0	73.2	72.4	71.9	72.7
6	87.8	81.7	88.5	84.7	61.4	63.5	84.1	79.1	86.1	88.3	67.1	62.7
7	54.1	46.3	73.6	67.1	64.6	64.5	53.7	46.7	65.3	64.8	64.9	66.3
8	66.9	65.3	92.8	89.1	95.6	101.6	70.4	65.3	86.9	83.7	89.8	94.4
9	72.5	66.9	64.9	56.8	74.4	68.1	64.2	63.2	60.2	55.1	75.8	69.8
10	68.6	63.1	63.7	59.2	73.0	71.4	69.9	62.1	61.5	66.2	76.2	76.6
Group Mean	75.7	70.0	77.1	73.0	71.6	72.1	75.5	68.8	73.5	72.6	72.0	71.7

* Group I, non-instructed; Group II, instructed; Group III, instructed-avoidance.

DISCUSSION

In examining possible explanations of this difference between the instructed and the

[4] Complete trial-by-trial data for each *S* may be secured from the American Documentation Institute, 1719 N Street, N.W., Washington 6, D. C., by ordering Document 3425, remitting $1.00, for 35 mm. microfilm, or $3.60 for 6 x 8 photocopies.

[5] In order to verify these observations, the means for the first five and last five trials within each group were compared by Wilcoxon's nonparametric test for paired items. The results support the picture given by Figure 2—only the instructed group's decrease in magnitude of the conditioned response during the course of extinction approaches significance ($.06 > p > .05$). The non-instructed group's response level remains high throughout extinction; that for the instructed-avoidance group remains low.

instructed-avoidance group, probably the first hypothesis to be entertained is that the very movement of *S* in pressing the key (avoidance response) results in an acceleration of the heart rate. This acceleration, in turn, counters the conditioned deceleration in heart rate and makes for an artifactual "pseudo-extinction." In order to test this hypothesis experimentally, it was necessary to use a control group of ten *S*s. This control group was given a basal series of 20 tone-alone trials, just as the experimental groups, except that no shock electrodes were placed on the hand. In the preliminary instruction, *S*s were told, "This is an experiment to determine your heart-rate reaction to tone." No reference to shock was made. Upon completion of the preconditioning series, the experimenter placed the key in front of the control *S*s (in the same position as for the experimental *S*s) and said, "From now on, I want you to press this key as soon as you hear the tone signal." This phase of the control group's performance corresponded to the extinction phase of the instructed-avoidance group. Pretone and posttone heart-rate records were taken throughout. The results indicated that there was no significant change in heart rate (acceleration or deceleration) for the control group during the test period (mean pretone heart rate was 73.5 b/m; mean posttone, 72.7 b/m).

The possible influence upon the heart rate of the motor response per se is not completely rejected by the above finding. It may be argued that the instructed-avoidance group pressed the key harder and faster than the control group, the former being more highly motivated. Data are not available to determine whether the instructed-avoidance group actually did press harder, but complete data are available to check upon the latency of the motor response. These data indicate that there is no significant difference in reaction time between the two groups; indeed, what difference exists reveals that the control group was the more rapid of the two (a median of 0.50 sec. for the control; 0.65 sec. for the experimental).

In view of the above findings, the rapid decrease in response strength of the instructed-avoidance group may be tentatively regarded as a valid extinction phenomenon, and not as an artifact of the motor response. Although prior verbal information itself facilitates extinction, as revealed by the data of Group II, a comparison of Group II with Group III indicates that a much more rapid extinction is engendered by the use of avoidance instructions plus an avoidance response. How much of this speedier extinction may be attributed to the avoidance instructions variable, however, and how much to the avoidance response itself cannot be completely determined from the present data.

SUMMARY

An experiment was performed to determine the effect of three different post-acquisition procedures upon the course of extinction of a conditioned heart-rate response in humans. The response was conditioned under a partial reinforcement schedule, with electric shock as the reinforcing agent. In a group extinguished without prior information (the non-instructed group), the conditioned response showed little tendency to decline in strength over 11 extinction trials. Following the first extinction trial, both an instructed and an instructed-avoidance group were told that they would no longer be shocked. In the latter instance, however, *S*s were told that the shock would be omitted only if they made a specific motor response (tapping a telegraph key) whenever the conditioned stimulus (tone) was presented. Although in the case of the instructed group the strength of the conditioned response declined progressively, the addition of the avoidance response for the instructed-avoidance group led to a much more rapid extinction.

REFERENCES

1. COOK, S. W., & HARRIS, R. E. The verbal conditioning of the galvanic skin response. *J. exp. Psychol.*, 1937, 21, 202–210.
2. HILGARD, E. R., & MARQUIS, D. G. *Conditioning and learning*. New York: D. Appleton-Century, 1940.
3. JENKINS, W. O., & STANLEY, J. B. Partial reinforcement: a review and critique. *Psychol. Bull.*, 1950, 47, 193–234.
4. NOTTERMAN, J. M., SCHOENFELD, W. N., & BERSH, P. J. Conditioned heart rate response in human beings during experimental anxiety. *J. comp. physiol. Psychol.*, 1952, 45, 1–8.
5. NOTTERMAN, J. M., SCHOENFELD, W. N., & BERSH, P. J. Partial reinforcement and conditioned heart rate response in human subjects. *Science*, 1952, 115, 77–79.
6. SCHOENFELD, W. N. An experimental approach to anxiety, escape, and avoidance behavior. In P. H. Hoch and J. Zubin (Eds.), *Anxiety*. New York: Grune and Stratton, 1950.
7. WILCOXON, F. Individual comparisons by ranking methods. *Biometric Bull.*, 1945, 1, 80–83.

Received November 29, 1951.

AN EXPERIMENTAL INVESTIGATION OF THE HYPNOTIC DREAM[1]

ANDERS SWEETLAND AND HERBERT QUAY
Florida State University

At first glance the hypnotically induced dream would seem to be an excellent method for the experimental exploration of the normal sleeping dream. As Luria has amply demonstrated (6), hypnosis can provide an *in vito* approach which obviates many of the controls necessary in *in vivo* experiments. The experimenter, using hypnosis, can frequently circumvent many of the problems of uncontrolled variables which plague clinical research. The method is consequently tempting. However, phenotypes must be presumed until genotypes can be demonstrated. And, until this, the hypnotic dream must be studied for itself and generalization beyond the data must be reserved.

The present paper is an exploratory study of the hypnotic dream per se. It is essentially an extension of the works of Klein (4) and Farber and Fisher (1), which were more demonstrative than experimental. It is an attempt to relate the response to two variables: the organismic and the stimulus. The questions asked are: how are reaction time, dream duration, recall, and dream alteration related to ego-involving stimuli; and does any relationship exist between response variables and organismic factors such as social-emotional adjustment and intelligence.

Procedure

Prior to the experiment, 16 psychologically naive college students were taught to dream under hypnosis. During practice they were encouraged to "dream the same as you do at night." The experiment was begun when subjects (*Ss*) could no longer distinguish between their hypnotic dreams and their normal sleeping dreams.

The *Ss* were randomly selected only to the extent that volunteers for such an experiment are random. Test results showed that ranges of intelligence and emotional adjustment were large, including the lowest and highest deciles.

The stimuli were 50 statements of an ego-involving nature: *You overhear somebody say that he wishes he had your personality. You are accused of being a social climber, etc.* These were obtained by having each member of a class in psychology list several items following a discussion of ego-involvement.

The items were edited, to make them generally applicable, and then given to 40 college students (20 males and 20 females) to rank on a seven-point scale. The ranking yielded 20 items paired in dispersion and distance from either side of a zero point. These 20 were paraphrased and ten neutral items added: *You have fish for Friday dinner,* etc. The result was 50 items fairly equally balanced around a neutral point. Immediately prior to experimentation, each *S* ranked this final set of stimuli on a seven-point scale.

To determine the effects of ego-involvement, the rankings were divided into three groups. Rankings one and two (markedly ego-enhancing) were called "plus." Rankings six and seven (markedly ego-assaulting) were called "minus," the remaining rankings "neutral." This division, while arbitrary, was supported by the fact that the extreme rankings tended to be stable: items thus ranked would be ranked similarly by all *Ss*. Although there was some shift of items in the neutral range, only rarely would a neutral item be ranked in the extremes.

Farber and Fisher (1) indicate that the characteristics of a dream may be influenced by a change in the people present. We attempted to control this by having the same people present for each *S*. For males this was either (or both) of the experimenters. Each female, however, was requested to bring a "chaperone." She was asked to bring someone towards whom she had no intense emotional feelings, either attraction or dislike. This generally was the girl's roommate. All experimentation was conducted in a soundproof room. Dreams were timed by having the *S* raise his hand when he began each dream and lower it on its completion.

After dreaming to the 50 items (an average of three one-hour sessions) the *Ss* were randomly divided into an experimental and a control group. The control group immediately repeated the 50 stimuli. The experimental group was given a discussion of the theory of dream symbolization including dream work, the censor, and manifest and latent content. The experimental group then repeated the 50 items, again keeping the same people present.

The results of this repetition are not included in the present experiment, which is restricted to quantitative findings. When possible, however, we have repeated the statistical computations on this second group of data as a check on the first. The data (obtained from the repetition) may be found in parentheses.

Results

Reaction Time and Dream Duration

The reaction times and duration of dreams were compared by the mean variation

[1] This study was made possible by a grant from the Research Council, Florida State University.

method (5, p. 67). The results show no reliable differences suggesting that ego-involvement has no influence on either. The data are summarized in Table 1. There is also no apparent relationship between dream duration and reaction time: rho=0.196±0.26.

TABLE 1

REACTION TIME AND DURATION OF DREAMS TO THREE TYPES OF EGO-INVOLVING STIMULI

	PLUS	NEUTRAL	MINUS
Reaction M_w	14.51 (15.81)	12.34 (13.07)	14.27 (14.55)
S^2_w	7.59 (12.97)	5.79 (8.75)	6.30 (10.57)
Duration M_w	13.71 (12.23)	13.85 (13.01)	15.14 (14.12)
S^2_w	5.99 (3.08)	4.49 (3.15)	6.70 (5.10)
	PN	PM	NM
Reaction t*	0.59 (0.59)	0.06 (0.26)	0.55 (0.34)
Duration t*	0.04 (0.24)	0.23 (0.66)	0.23 (0.26)

* For 15 *df*, *t* must be at least 2.131 to be reliable at the 5 per cent level. The numbers in parentheses are obtained from the repeated experiment. See text.

The data in Table 1 tend to conceal a marked personal equation. Inspection of the protocols showed that each *S* had his own speed of reaction and dream duration, generally consistent within a few seconds. This personal equation was also evidenced in the narratives. Music majors tended to dream about music; the journalism student dreamed about newswriting; the camera fan dreamed about photography. One male *S* dreamed repeatedly of nude females. We tactfully made no inquiry.

The data may give some indication of the answer to the classical question: how long does a dream last. The length of the average hypnotic dream is ten to fifteen seconds. The range of individual dream duration is large: one second to five minutes.

The Relation of Alteration to Stimulus Tone

The main consideration of the experiment was the behavior of the "dream work": the alteration (symbolization) of the stimulus material. A dream was classified as being either an alteration or a repetition of the stimulus. Dreams which were essentially a reduplication or a continuation of the stimulus (or, frequently, both) were classified as "repeat." Dreams which contained definite elements of typical dream distortion were classified as "alteration." The term *alteration* was used to distinguish the phenomenon from symbolization as conventionally used.

Two examples of "repeat" to the stimulus "You are accused of being a social climber," are: "I was walking down the hall when I heard these two girls say this. I was furious and asked them what they meant. . . ." "When I heard this I made up my mind that I would try to find out why they said this and do something about it."

Examples of "alteration" (same stimulus) are: "I was climbing up a ladder. There were people all around. They seemed to be mad at me, but I kept climbing." "It seems that my roommate said this. I don't exactly remember, but I remember that the room was all in white. It had a white bed, and a white dresser, and white walls. This was very strange to me because I knew it wasn't like that. Then I ran out of the room."

The dreams were scored independently by the two experimenters. Most samples of 50 dreams showed 100 per cent agreement. Lowest agreement was 92 per cent. An occasional dream would be classified as "repeat" by one scorer and "alteration" by the other. If discussion did not resolve the disagreement, the classification was assigned randomly. This happened in less than 1 per cent of the judgments.

Because of the close agreement between judges, no reliability coefficients were computed. It is believed that the agreement well exceeds that necessary for making group comparisons.

The first consideration was whether the distribution of alteration in each of the stimulus categories (plus, minus, and neutral) was the result of stimulus tone or chance. To determine this a chi square was obtained between the observed and expected frequencies. The expected frequency in this, and subsequent cases, was obtained from the proportional number of dreams in each category. In a strict sense, our use of chi square may not be defensible as there is some question of cell independence. It is felt, however, that

the conclusions are not influenced by the choice of statistic.

A chi square of 5.21 (p between 5 and 10 per cent) was obtained. This is probably a chance variation as a similar chi square from the second group of data had a probability at the 60 per cent level of confidence. The results indicate no tendency to alter any particular type of stimulus used in this experiment.

The Relation of Dream Alteration to Emotional Adjustment and Intelligence

The unexpected findings of the experiment were the relations between dream alteration and personality adjustment. Personality adjustment was measured by two tests: The Social Security-Insecurity Test (SI) by Maslow, *et al.* (7) and the Minnesota Multiphasic Personality Inventory (MMPI) of McKinley and Hathaway (2).

The SI is a 75-item inventory, designed to measure adjustment in the area of social security. The MMPI is a 560-item inventory. It has separate scales for Hypochondriasis (*Hs*), *Depression* (*D*), Hysteria (*Hy*), Psychopathic Deviate (*Pd*), Masculinity-femi*nity* (*Mf*), Paranoia (*Pa*), Psychasthenia (*Pt*), Schizophrenia (*Sc*), and Hypomania (*Ma*). It also has several validating scales. Two of these were used, *F* and *K*. The *F* scale, originally designed to check sorting and clerical errors, has been found to correlate with adjustment. The *K* scale was designed to control the effects of the attitude of the person taking the test.

We had originally designed the experiment to include only *S*s of average social-emotional adjustment. To check this, each *S* was given the SI. On scoring several of these it became apparent that there was a relationship between the amount of dream alteration and the adjustment scores. To check this further, the MMPI was given. The relation between the number of alterations and the scores on the SI and the MMPI may be found in Table 2.

The results given in Table 2 indicate a definite relation between hypnotic dream alteration and test scores. This relationship may be stated: the better the adjustment of the individual, the greater the amount of alteration he produces. The correlation between SI scores and alteration is reliable at the 5 per cent level (Fisher's z). Although none of the correlations with the MMPI (except that with the *K* scale) is reliable, the chance of obtaining eight negative correlations, assuming zero correlation, is 1 in 256.

TABLE 2

CORRELATIONS BETWEEN DREAM ALTERATION AND ADJUSTMENT SCORES

SI	.574	(.564)
MMPI		
F	−.028	(−.024)
K	.720	(.571)
Hs	−.004	(.003)
D	−.235	(−.338)
Hy	−.211	(−.236)
Pa	−.341	(−.522)
Pt	−.288	(−.243)
Sc	−.088	(−.099)
Ma	−.151	(−.079)

To check the relationship further, correlations were computed between intelligence and alteration and intelligence and adjustment. The measure of intelligence was the American Council on Education Psychological Examination (ACE). This test has three subscales: Quantitative (Q), Linguistic (L), and a Total Score (T). We found small correlations between the ACE and SI scores. This was partialled out. The original correlations and the partialled correlations may be found in Table 3.

TABLE 3

CORRELATION BETWEEN ADJUSTMENT SCORES (SI), INTELLIGENCE SCORES (Q, L, AND T) AND ALTERATION (Sy) *

PRODUCT MOMENT CORRELATIONS

r	Q	L	T
SI	.327	.307	.343
Sy	.367	.339	.382
r_{SIsy}		.574	

PARTIAL CORRELATIONS

SySI.Q	.515
SySI.L	.520
SySI.T	.510
SyQ.SI	.231
SyL.SI	.334
SyT.SI	.240

* Correlations exceeding .482 are reliable at the 5 per cent level.

The results shown in Table 3 support the thesis that there is a real relationship between dream alteration and adjustment. The data also suggest a slight relationship between alteration and intelligence.

Rejections

Occasionally no dream followed the presentation of the stimulus. We have called this a rejection. There was no stimulus that all *Ss* rejected; some *Ss* rejected none. Feeling that rejections might have some significance, we determined the number of rejections from plus, minus, and neutral stimuli and computed the chi square between this and the expected frequency. The expected frequency was determined from the total number of dreams in each category.

The chi square obtained was 2.63. For 2 *df*, this is between the 20 and 10 per cent levels. It indicates that this variation can be attributable to chance factors. This is further supported by the chi square obtained from the second group, which was between the 70 and 50 per cent levels. Evidently rejections are not related to ego-involvement.

Inquiry indicated that some of the rejections resulted when *Ss* felt the stimulus inapplicable or implausible. The subject who didn't smoke, for example, tended to reject the stimulus "You find that you are out of cigarettes and purchase a package immediately."

Memory of Dreams

Two aspects were considered. the relationship of recall to stimulus tone and the relationship of recall to primacy and recency in an hour's dreams.

We awakened each *S* at the end of each experimental session and asked him to recall as many dreams as he could. We then compared the number of dreams recalled in each of the categories (plus, minus, and neutral) with the expected number (determined from the proportional number of dreams in each category). A chi square of 2.61 (*p*, 25 per cent) was obtained. Recall is evidently not related to stimulus tone. It should be remembered that this is based on stimulus tone, not response tone.

The effect of primacy and recency on recall was determined by dividing each session's dreams into chronological fifths. The frequency of remembered dreams was tabulated for these fifths. A chi square was computed between this frequency and the expected frequency, assuming that recall would be equally distributed. The chi square of 21.48 was significant at the 1 per cent level (4 *df*). Inspection of the computations showed that over two-thirds of the chi square (14.51) came from the large frequency in the last quintile.

To check this, the last quintile was removed and the chi square computed for the remaining data. This chi square (4.35) was reliable at the 25 per cent level (3 *df*). The indication is that the last dreams tend to be remembered most.

DISCUSSION

The problem most interesting, and as yet unresolved, is the relationship of the hypnotic dream to the normal sleeping dream. Sirna (8), in a study using the electroencephalogram, concluded that the hypnotic dream is psychologically, but not physiologically, the same. His *Ss* reported that their hypnotic dreams were indistinguishable from the sleeping dreams. He based his physiological differences on the difference in brain waves in sleep and in the hypnotic state. Farber and Fisher (1) similarly report that their *Ss* could not distinguish between the hypnotic and the sleeping dream.

Klein (4) reports that his *Ss* unanimously agree that the hypnotic dreams can legitimately be called dreams. The only differences seemed to arise from the fact that the dreams were produced in the hypnotic state. He concludes that discrepancies were the result of the setting and not the dream process.

Our *Ss* did not begin experimentation until they could distinguish no appreciable difference between the sleeping dream and the hypnotic dream. Similarly to Klein, we found that differences were a confusion of the situation and not of the dream per se. Further research on the problem is indicated before any definite statement can be made. We are, however, impressed by the similarity and are endeavoring research in this area.

Some observations did not lend themselves to quantification. Because of their theoretical implications we have included them in the following.

Klein, using sensory stimuli (touch, smell, sound, etc.) found that some *Ss* had the same or similar dreams when the stimulus was repeated. Our stimuli differ: they are more conceptual in nature. On repeating them, we found very few repeated dreams: less than five in 800 repetitions. This is despite the fact that the stimuli were repeated in an identical setting and identical sequence. It would seem that the occurrence of the repeated hypnotic dream is more closely related to "peripheral" than to "central" stimuli.

It was also observed that some *Ss* tended to "pretty up" the dream content. Although their dreams were not altered, ego-assaulting stimuli were made to "come out all right in the end." It is our plan to quantify this in later research.

Some *Ss* persistently yielded dreams which had no apparent relationship to the stimulus, even when the stimulus was available for comparison. There were three protocols in which it was impossible to locate a dream from the stimulus by any method other than counting down the list.

An illustration of this remoteness of relationship between the dream and its stimulus is given in the following response: "I saw two men walking down the street talking to each other. One was walking forward and the other backward. They were gesticulating wildly and I remember they walked past a pawnshop." The stimulus was: "You have fish for Friday dinner." Other dreams were even more remotely associated with the stimulus, involving geometrical light formations, abstract designs which flowed into shapes and again into abstractions, and personifications of the dreamer into lines of poetry, notes from a song, etc.

The dreams which were remote from the stimuli were invariably dreamed by *Ss* having excellent adjustment scores on the personality tests. As adjustment scores dropped, the dreams became more and more reduplications of the stimuli, the poorest adjusted *S* showing the most trite reduplication.

The hypnotic dream seems to be closely related to organismic factors and remotely (if at all) related to stimulus factors. We should like to caution against generalizing beyond the materials used in this experiment as the use of more intense stimuli might yield different results. Although we are impressed with the similarity of the hypnotic dream to the sleeping dream, it is our feeling that interpretations made from one to the other should be held in abeyance until further research is available.

We feel that the hypnotic dream is an imaginative phenomenon and, as such, is sensitive to emotional adjustment. This interpretation tends to classify lack of alteration more in the category of the psychological deficit as described by Hunt (3). This idea is advanced tentatively. Again, further research is indicated.

SUMMARY

Sixteen psychologically naive college students were taught to dream under hypnosis. Fifty stimuli of an ego-involving nature were twice administered. It was found that:

1. Ego-involvement showed no effect on reaction time, dream duration, dream alteration ("symbolization") and stimulus rejection.

2. Dream alteration is closely related to emotional adjustment; the best adjusted subjected symbolizes most.

3. Recall is a function of recency and is independent of stimulus tone.

4. Repeated dreams occur rarely when "conceptual" (as opposed to sensory) stimuli are used.

REFERENCES

1. FARBER, L. H., & FISHER, C. An experimental approach to dream psychology through the use of hypnosis. *Psychoanal. Quart.*, 1943, **12**, 202–216.
2. HATHAWAY, S. R., & MCKINLEY, J. C. *Manual for the Minnesota Multiphasic Personality Inventory*. New York: Psychological Corporation, 1943.
3. HUNT, J. MCV., & COFER, C. N. Psychological deficit. In J. McV. Hunt (Ed.), *Personality and the behavior disorders*. New York: Ronald, 1944. Pp. 971–1032.
4. KLEIN, D. B. The experimental production of dreams during hypnosis. *Univer. of Texas Bull.*, 1930, No. 3009.
5. LINDQUIST, E. F. *Statistical analysis in educational research*. Boston: Houghton Mifflin, 1940.
6. LURIA, A. R. *The nature of human conflict*. New York: Liveright, 1932.
7. MASLOW, A. H., HIRSCH, E., STEIN, MARCELLA, & HONIGMAN, IRMA. A clinically derived test for measuring psychological security-insecurity. *J. gen. Psychol.*, 1945, **33**, 21–41.
8. SIRNA, A. J. An electroencephalographic study of the hypnotic dream. *J. Psychol.*, 1945, **20**, 109–113.

Received December 1, 1951.

AGGRESSION IN PROBLEM CHILDREN AND NORMALS AS EVALUATED BY THE ROSENZWEIG P-F STUDY

SAUL ROSENZWEIG AND LOUISE ROSENZWEIG

Community Child Guidance Clinic
Washington University
St. Louis, Missouri

THE golden days of childhood, as they appear in the conventional reminiscence of adults, are more consistent with fantasy than with fact. The myths of all races similarly portray the childhood of man as a paradise. Religions complement the illusion by projecting paradise into a future life. The most direct conclusion is, perhaps, that adult life, here and now, is characterized by the frustration and unhappiness of "responsibilities"; the irresponsible years of childhood are thus conjured up with yearning to ease the pain of present burdens and to serve as a springboard for a future paradise without care. But if the child himself is consulted, the testimony is, unfortunately, more equivocal. The child is probably no happier than the adult. The happiness of the one is too closely interwoven with the happiness of the other for much difference to be possible. And, while the adult looks back to his childhood longingly, the child correspondingly anticipates the time when he will be "grown up" and become free of the parental control which constantly frustrates him. This interplay of frustration in child and adult is nowhere more clearly apparent than in the problems brought to the Child Guidance Clinic.

In the foregoing setting the modest purpose of the present communication is twofold: (*a*) to present tentative norms for a group of child guidance patients on the Rosenzweig Picture-Frustration (P-F) Study; (*b*) to compare these norms with those previously reported for normal children (3). It was believed that significant differences between the two groups, if found, would throw light on the dynamics of aggression in childhood maladjustment.

The population on which the present norms are based consisted of 162 children who were seen at the Community Child Guidance Clinic of Washintgon University during the period between January 1, 1949 and April 30, 1951. They were referred for a variety of emotional disturbances in which some deviation from accepted social functioning was always evident. All patients between the ages of four and thirteen who were seen at the Clinic during the above-mentioned period and were administered the P-F Study have been included. The Study was always administered individually. Of the total group, 103 were males and 59, females. The preponderance of males reflects the similar difference in referrals for the two sexes. Since sex differences were not previously found to be very significant for normal children, such differences were not examined in the present analysis.

In Table 1 are given the patient norms for the various P-F categories and for the Group Conformity Rating (GCR).[1] Table 2 presents similar findings for the various P-F factors; Table 3, the data on the so-called superego factors and patterns (E, I and their various combinations).[2]

Comparison of Problem with Normal Children

Comparison will be oriented with respect to (*a*) over-all developmental trends throughout the age range studied in the two populations and (*b*) cross comparison between the two groups at each year level considered.

As regards the percentages of the various scoring categories and GCR, similar developmental trends are found in both normal and patient groups for Direction of Aggression. In particular, E decreases consistently with age while I and M tend to increase. GCR increases with age up to the 12–13 year level for both populations but at this point drops slightly in both. For Type of Reaction the

[1] The P-F categories, factors, and patterns are defined and described elsewhere (1, 2, 3).

[2] The statistical assistance of Esther Lee Mirmow and Leo Burns is gratefully acknowledged.

TABLE 1

MEANS AND STANDARD DEVIATIONS OF SCORING CATEGORIES AND GCR BY PERCENTAGE FOR CHILD GUIDANCE PATIENTS

YEAR LEVEL	E	I	M	O-D	E-D	N-P	GCR
4–5 (N=12)							
Mean	69.33	12.25	18.42	16.58	59.92	23.50	46.42
SD	16.57	9.75	12.52	10.19	14.61	11.71	9.66
6–7 (N=38)							
Mean	63.16	18.42	18.26	16.21	56.50	27.34	56.76
SD	16.24	9.00	12.88	6.87	14.86	11.61	11.41
8–9 (N=47)							
Mean	54.19	20.85	24.83	15.61	58.76	25.47	60.06
SD	18.10	8.94	13.23	7.04	13.30	13.93	11.89
10–11 (N=36)							
Mean	47.47	22.53	30.08	18.69	53.00	28.30	61.72
SD	11.75	6.26	9.73	6.48	9.08	9.34	10.20
12–13 (N=29)							
Mean	43.93	25.14	30.79	19.69	49.86	30.45	56.41
SD	20.06	9.55	14.30	7.58	12.90	13.91	10.90
Composite (N=162)							
Mean	54.09	20.78	25.05	17.24	55.44	27.28	57.99
SD	18.55	9.23	13.58	7.43	13.38	12.50	11.75

normals show no consistent developmental trends but the patients appear progressively to decrease in E-D and to increase in N-P. These latter indications are, however, small and not maintained throughout.

The cross comparison at each age for the two populations in respect to Direction of Aggression shows consistent differences as follows: (*a*) E percentage is consistently higher for the patients than for the normals at every level (at 4–5, 6–7, and 8–9, $p<.05$). (*b*) The patients are lower than the normals at each age level on I (at 4–5, 6–7 and 12–13, $p<.05$). (*c*) Less consistent difference is found for M percentage. There is, however, some tendency for the patients to be lower than the normals at the three younger levels ($p<.05$ for 8–9 only). (*d*) GCR is lower for patients than for normals at every age. (Statistical significance is, however, found only for 12–13, $p<.01$.) (*e*) Comparisons with respect to Type of Reaction show no consistent differences for N-P but demonstrate a lower E-D percentage for the patient population at

TABLE 2

MEANS AND STANDARD DEVIATIONS OF SCORING FACTORS BY FREQUENCY

YEAR LEVEL	E′	E	e	I′	I	i	M′	M	m
4–5 (N=12)									
Mean	1.29	11.25	3.71	.54	1.71	.58	2.04	1.04	1.17
SD	1.35	5.00	1.79	.69	1.90	.76	1.62	1.48	1.38
6–7 (N=38)									
Mean	1.54	8.87	4.56	.76	2.96	.66	1.54	1.54	1.28
SD	1.29	4.12	2.45	.62	1.75	.90	1.00	1.62	1.41
8–9 (N=47)									
Mean	1.35	7.91	3.66	.61	3.64	.74	1.78	2.47	1.69
SD	1.30	4.39	2.58	.67	2.07	.95	1.34	1.78	1.26
10–11 (N=36)									
Mean	1.71	6.03	3.60	.65	3.97	.75	2.12	2.67	2.39
SD	1.05	2.22	2.12	.67	1.36	.97	1.48	1.47	1.30
12–13 (N=29)									
Mean	1.36	5.33	3.81	.60	3.98	1.43	2.71	2.62	2.05
SD	1.40	4.04	2.47	.61	1.91	1.88	1.68	1.70	1.48

TABLE 3

MEANS AND STANDARD DEVIATIONS OF SUPEREGO FACTORS AND PATTERNS BY PERCENTAGE

YEAR LEVEL	E	I	E+I	E—E	I—I	M+I
4–5 (N=12)						
Mean	10.33	1.92	12.33	37.50	5.42	20.33
SD	10.05	2.72	9.21	15.28	7.32	13.37
6–7 (N=38)						
Mean	12.18	5.37	17.66	25.32	6.95	23.89
SD	7.70	5.49	7.98	15.17	5.74	13.41
8–9 (N=47)						
Mean	10.06	6.34	16.64	23.08	8.55	31.40
SD	8.56	4.81	8.77	14.26	8.19	13.69
10–11 (N=36)						
Mean	9.92	7.03	17.39	14.94	9.25	37.30
SD	5.05	4.92	6.08	7.78	5.48	11.24
12–13 (N=29)						
Mean	8.14	8.55	17.00	13.96	7.86	39.38
SD	5.34	6.09	6.59	13.13	5.78	15.09

year levels 10–11 and 12–13 ($p<.05$ for both) and a higher O-D percentage at 10–11 ($p<.01$) and 12–13 ($p<.05$).

The foregoing findings for the various categories concerned with Direction of Aggression and for the GCR are graphically portrayed in the accompanying bar diagrams (Fig. 1).

In the further presentation of the results for the several P-F factors and for the superego patterns, the present item populations are too small and the distributions too skewed to permit the use of statistical tests of significance for found differences. The findings to be reported are regarded as highly tentative and should be so construed by the reader. It will be noted that most of the conclusions in the final section depend, in largest part, only upon the above-given results for the P-F categories and GCR.

For the factors, developmental trends are not found consistently in either population except for E, which in both groups steadily decreases with age, and I, which in both shows some tendency to increase. Additional indications, for the patients only, appear with respect to three factors—i, M, and m—all of which tend to increase with age.

Comparison of the two populations at the several age levels indicates e to be higher for the patients at every age; I, lower for them at every level except 8–9; M, slightly lower at the first three age levels; m, lower at all levels.

The developmental trends for superego factors and patterns are similar for both populations: E decreases, albeit inconsistently; I rises consistently. In like manner, E — E

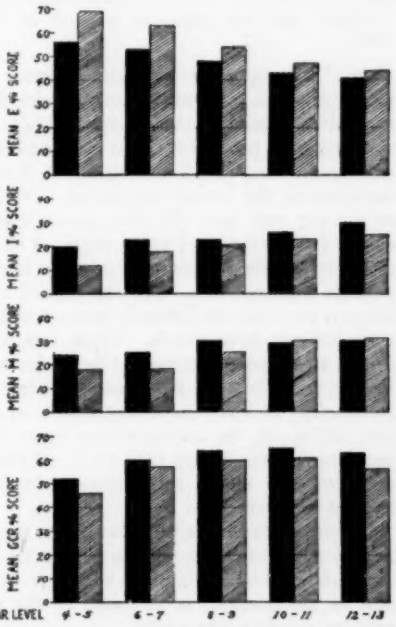

FIG. 1. COMPARISON OF CHILD GUIDANCE PATIENTS WITH NORMALS

Black=Normal; Diagonal=Child Guidance.

decreases and M + $\underline{I}$ increases with age in both populations. The tendency for I — $\underline{I}$ to increase is slightly better maintained in the patient than in the normal group.

The comparison of normals and patients for the superego variables adds but little to the foregoing. The data suggest, however, that $\underline{E}$ is slightly higher and $\underline{I}$ similarly lower for the patients than for the normals at several of the age levels; I — $\underline{I}$ is at every age lower; M + $\underline{I}$ is lower for the first three levels.

In Table 4 is shown the incidence of

TABLE 4

INCIDENCE OF TRENDS FOR CHILD GUIDANCE PATIENTS AT EACH AGE LEVEL

AGE	*N* *	*N* WITH TRENDS	% WITH TRENDS
4–5	12	7	58
6–7	38	27	71
8–9	47	34	72
10–11	36	34	94
12–13	28	17	59

* Trends could not reliably be computed for one subject.

trends[3] among child guidance patients at each age level. Corresponding results for normals have been previously reported (4). A comparison of patients with normals indicates that trends are found in somewhat fewer members of the former population. While at the 4–5 year level the two groups agree closely (58 per cent and 57 per cent), at three of the other four levels trends occur in fewer patients than normals. The exceptional age range is 10–11 where distinctly fewer normals than patients have trends. These tentative results are supported somewhat by the further finding that, while the mean number of trends for normal children has been shown to be 2.07 (*SD* 1.19), the corresponding figure for problem children is 1.77 (*SD* 1.61). (Difference between means significant at the .05 level of confidence.)

It will be noted that for most of the mean figures reported in the four preceding tables the variability of the patients is greater than that of the normals, as indicated by the respective standard deviations.

[3] Attention is called to the technical meaning of *trends* as employed in this part of the report. (See 1 and 4.)

DISCUSSION

The preceding results with respect to P-F categories, factors, and patterns, considered *developmentally* for both normal and problem children through the several age groups, warrant the following tentative conclusions: (*a*) "Aggression" (scored in the P-F Study as extrapunitive) is more completely expressed in the earlier than in the later years of childhood. There is a gradually diminished direct expression of aggression from the 4–5 year level, step by step, through the 12–13 year level. (*b*) *Pari passu* with this developmental change is the progressive increase of indirect or deflected expressions of hostility (the progressive increase of intropunitive and impunitive responses above noted). (*c*) The increasing socialization of the child is reflected in the continuously rising Group Conformity Rating and in the similar changes in superego patterns (especially E — $\underline{E}$ and M + $\underline{I}$). (*d*) These findings (*a* through *c*) lend support to the psychoanalytic formulations regarding the growth of the so-called superego. They are also, of course, consistent with other observations concerning the child's developing self-control during the period from four years to puberty.

The cross comparison between problem and normal children at each age level indicates that the maladjusted child is distinctly less able to control aggressive responses in frustrating situations than is the well-adjusted child. Whether the data warrant the further conclusion that such problem children have less frustration tolerance is perhaps more open to question. An alternative and more conservative formulation would be that they have less fully acquired the patterns of socially acceptable reaction to frustration—or that they find it more difficult to participate in such social learning. The consistently lower adaptability of the patients is indicated in the data (frequency of m); while the normals show practically no change in this respect from one age level to the next, the patients give evidence of attaining the same degree of adaptability gradually and belatedly. The greater variability of problem as compared with normal children must be noted as an important qualification in the understanding of individual problem children who

may actually sometimes be over-conforming or excessively self-controlling.

The findings with respect to trends point again to the relative immaturity of the problem children. In a previous investigation of trends among normal children, it was shown that there is a statistically significant difference for frequency of subjects with trends at ages 4–5 as compared to all later levels. The results were interpreted "as reflecting the process of socialization during childhood with an attendant suppression of infantile aggressiveness and concomitant arousal of conflict in reacting to frustration" (4, p. 313). In these same terms the results for the present child patients, who give evidence of having fewer trends than normals, may be taken to mean that these subjects remain less socialized than their better-adjusted peers well beyond the fifth year, and hence have less self-critical conflict in expressing aggression. This interpretation is, of course, in agreement with the results for the various P-F categories and the GCR.

It may in conclusion be observed that the comparison of problem children with normals, as presented above, affords a modicum of validation for the Children's Form of the Picture-Frustration Study; i.e., the differences between the two populations are, in general, consistent with what one might expect from the fact that the problem children have "problems" and are, therefore, presumably more frustrated.

REFERENCES

1. ROSENZWEIG, S. The picture-association method and its application in a study of reactions to frustration. *J. Pers.*, 1946, **14**, 3–23.
2. ROSENZWEIG, S., FLEMING, EDITH E., & CLARKE, HELEN J. Revised scoring manual for the Rosenzweig Picture-Frustration Study. *J. Psychol.*, 1947, **24**, 165–208.
3. ROSENZWEIG, S., FLEMING, EDITH E., & ROSENZWEIG, LOUISE. The Children's Form of the Rosenzweig Picture-Frustration Study. *J. Psychol.*, 1948, **26**, 141–191.
4. ROSENZWEIG, S., & MIRMOW, ESTHER L. The validation of trends in the Children's Form of the Rosenzweig Picture-Frustration Study. *J. Pers.*, 1950, **18**, 306–314.

Received December 24, 1951.

INFILTRATION AND THE ATTITUDES OF WHITE AND NEGRO PARENTS AND CHILDREN [1]

CHARLES BIRD, ELIO D. MONACHESI, AND HARVEY BURDICK
University of Minnesota

APPROXIMATELY three years ago, at the inception of the Laboratory for Research in Social Relations, University of Minnesota, one of the pilot studies selected was the determination of white and Negro interactions within two adjacent areas of the city of Minneapolis. There were several reasons for undertaking an investigation of this kind. Within the past decade public-minded citizens of Minneapolis have made serious attempts to reduce discrimination against Negro people; moreover, they have made it possible for Negroes to obtain employment in industrial, mercantile, and other establishments to an extent not visibly noticeable before. Since Negro families have lived within the city for many decades, their acceptance or rejection by white families, it was believed, would yield information about an important aspect of social responsibility. When the study was first planned, it was considered possible to institute techniques for measuring the reduction of group tensions if the data gained revealed similarity between the residents of the two adjacent districts. Up to this time, the collecting and collating of data have delayed the use of small group techniques and intensive personal interviews with selected residents from the two areas.

Too much stress cannot be put upon the fact that the residents participating in this study have not, as communities, manifested open antagonism to Negroes. The data that follow [2] reveal considerable prejudice expressed by individuals against Negroes but they do not imply a public expression of aggressive reactions. The respondents were markedly cooperative and genuinely interested in knowing the results of the inquiry. Of 146 white families selected for interviewing, only one refused to cooperate. The nature of the replies to several types of interviewing schedules invites confidence in the sincerity of the respondents.

METHOD

In order to preserve anonymity, the subjects of this study will be designated as residents of District X and District Y. Although these two districts are contiguous, District X is physically less desirable because its homes are older and closer to industrial and business establishments. The residents of both areas generally would be rated as belonging economically to the middle class and lower-middle class.

The major focus of this investigation into racial attitudes was the family, or, at least, the father, mother, and one or more children enrolled in the third, fourth, or fifth grade of either of two public schools in the adjacent districts. Through the cooperation of the assistant superintendent of schools, the principals of the two schools, Parent-Teacher Associations in the two districts, and a committee of Negro citizens, we were given access to all children enrolled in the above named grades.[3] All of the procedures we planned to use were explained to, and reviewed carefully by, representatives of these various groups. Even though modifications in measuring instruments had to be made to meet with general approval, without the cooperation accorded, this inquiry could not have been undertaken. It was the seal of approval of these groups that facilitated access to adult respondents by pairs of trained interviewers whose functions were to estimate the status of the home and to secure answers to a questionnaire from the mother and father separately and without knowledge of each other's replies.

The initial step of the inquiry was taken when 393 white boys and girls, and 31 Negro boys and girls, in the third, fourth, and fifth grades of the two schools answered an Attitude Scale, a Social Distance Scale, and a Home Index. For a detailed account of these scales the reader's attention is called to articles (5) and (4) of the bibliography. The Attitude Scale consisted of 26 sentences, such as, "They work hard," "I would like to be in a club or on a team with them," or, "I would like

[1] This is one of a series of studies on the broad problem of social responsibility being conducted in the Laboratory for Research in Social Relations, University of Minnesota, under a grant from the Carnegie Corporation.

[2] Not all of the data will be presented in this article. Three additional papers, in preparation, deal with other aspects of racial prejudice. These additional papers are: II. Methods for validating a questionnaire concerning white-Negro relationships by means of voluntary statements; III. The effect of parental discouragement of play activities upon the attitudes of white children toward Negroes; and IV. Prejudice and discontent.

[3] We take this opportunity to thank those who generously aided us even though their names cannot be used if the anonymity of the respondents is to be respected.

to go on a picnic with them." As used originally (5, pp. 85–86) the third person pronoun referred to Negroes, but in this investigation the children found the sentences tied to five ethnic or racial groups, namely, Germans, Jews, Indians, Japanese, and Negroes. The sentences were intended to suggest the willingness of children to engage in activities with members of the groups that were named, as well as to indicate how closely the opinions of children correspond with conventional stereotypes about racial and ethnic groups. The children were also asked to answer *Yes* or *No* to a list of nine questions comprising a Social Distance Scale, and again the same five groups were named as the focus of the ratings. For example, the children were confronted with statements such as, "I would let them visit our country," and, representing the other extreme of social distance, "I would be willing to marry one of them when I grow up." [4]

One other schedule was used in the preliminary survey of all the children in the three school grades. The teachers rated each child by means of the Haggerty-Olson-Wickman Behavior Rating Schedules.[5] We desired to know whether Negro children would be regarded by teachers as manifesting more behavior problems than white children.

Since our major objective was the study of the acceptance or rejection of Negroes by white parents and their children, by means of the scores on the Attitude and Social Distance Scales, a group of 152 children was selected from the 393 children for further investigation. Since relatively few Negro children were enrolled in the three grades, all of them, or 31 boys and girls, were also selected for further study. The parents of the 152 white children, 144 mothers and 128 fathers, and the parents of the Negro children, 24 mothers and 19 fathers, later became our adult sample. Several of these families were broken by death, divorce, or separation, so that a family was considered complete whenever both parents were interviewed or the only available parent was interviewed.

Following their selection to represent the population of the three grades in each of the two public schools, the children were interviewed individually and asked to answer 41 main questions and a smaller number of supplementary questions in a schedule called "Intergroup Understanding, Questionnaire for All Children." This questionnaire was especially prepared by members of the research team and subjected to criticism by the trained interviewers (University of Minnesota graduate students) who had pretested it in areas of the city sufficiently far removed from the districts in which the two schools were located. Not all of the questions were designed to plumb preferences for or prejudices against white and Negro children, since some questions were used as "shock absorbers" and others were intended to yield general information.

The questions selected for scoring, or to obtain a measure of attitudes, always named five racial or cultural groups, namely, Japanese, Negroes, whites, Jews, and Indians. To overcome any possibility that the ordering of these five groups would distort answers, the interviewers were provided with cards that permitted a changing order from question to question. The following questions were scored either +1 or −1, depending upon the implication of the reply with respect to preference for or prejudice against white and Negro children:

1. Do you play with boys (girls)[6] from any of these groups during school recess?
2. Do you play with boys (girls) from any of these groups when school is out?
3. If you were to pick a boy (or girl) for a side in a game, which group would you pick one from? All are just as good players. Which one next? And next? Any other?
4. You're choosing up sides. You can pick five children. The Negro (white, if Negro child is interviewed) children are all better players than any other children. From which group or groups would you pick your side?
5. Which group of children do you think are most likely to cheat in school?
6. Who do you think are most likely to play hookie or skip from school?
7. Do you mind having any of these families living near you?
8. Which do you think are the nicest families?
9. Are there any of these children who get into trouble more than others with your teachers?
10. Who do you think gets into trouble most with other boys and girls?

Since each answer was scored either +1 or −1, 20 points were added to each score to avoid the use of negative numbers. The possible range of scores, therefore, was from 10 to 30. Actually the obtained scores had a range from 12 to 30. The mean scores of white and Negro children on the questionnaire will be found in Table 3.

Another set of scores, indicating the intelligence levels of 144 of the 152 white children and 30 of the 31 Negro children, was made available by the school administrators.

A separate questionnaire was prepared for the parents of the children. It was also pretested by pilot studies. Trained interviewers administered the questionnaire in a face-to-face situation. All of the interviewing was done in the homes of the respondents and always with only one parent in the room. Although our main interest centers in the questions that were scored as indices of attitude, the kind of personal data obtained from a set of 13 preliminary questions may be mentioned. We inquired into such matters as length of residence in Minneapolis, other places of residence for periods longer than six months, length of residence in the present neighborhood, ownership of home, whether a renter would consider buying a home in the neighborhood, the last school attended and the last grade completed, the kind of occupation and all

[4] The authors of these scales were H. G. Gough, D. B. Harris, W. E. Martin, and M. Edwards, all of whom cooperated in administering them.

[5] Published by World Book Company, Chicago, 1930.

[6] Boys were asked these questions about boys; girls about girls.

other full-time jobs held, including which job was liked best and which job the respondent would really like to be doing, church affiliation including membership and frequency of attendance, and the number of children in the family together with their names, ages, and sex. There were other questions that aimed to tap conventional stereotypes of Negroes but that were not used to measure prejudice. These questions are presented in Table 2 together with the percentages of white respondents expressing differences in attitude and belief.

The questions that contributed to the total score of the interview schedule were:

1. Which of these groups (Japanese, Jews, Greeks, Negroes, Indians, native-born whites)[7] would you not like to have live in your neighborhood?
2. If you treated the members of these groups as good friends, would you get into trouble with your neighbors?
3. If stores in your neighborhood were owned and run by members of these groups, would you stay away from stores owned by any of them?
4. If you were invited into the homes of the groups named here, which invitations would you accept?
5. If you had an opportunity to drop in to see and chat with any of the following groups, which ones would you call upon?
6. Which of these groups are usually cleaner or neater in their homes?
7. Would you think of moving away from this neighborhood if any of these groups should move in?
8. Would you prefer to live in a neighborhood where only one of these groups lives?
9. Are there any of these groups of boys and girls you would not want your children to play or go with?
10. Would you let your boy or girl join an organization in which these groups belonged?
11. Have you discouraged any of your children from playing with or going with children in any of these groups?

Each of the above questions was scored either +1 or −1, depending upon the position assigned to Negroes in the first three or the last three ranks of the six groups named by the interviewer, or upon a Yes or No answer in the case of question 11. One other question, which is not listed above, occasionally contributed to the total score. Each respondent was asked: "Do you like or dislike living in this neighborhood? Why?" If a parent gave as a reason for disliking the neighborhood the presence of Negro residents, his response was assigned a score −2. Since no white respondent named Negroes as having a positive influence upon his liking of the neighborhood, a plus score could not be assigned to the question. The scoring scheme made possible, therefore, a range of scores from +11 to −13, but to avoid negative numbers 20 points were added to each score. Actually the obtained range of scores was identical with the possible range, so that with the 20 points added, the range was +7 to +31.

[7] These groups were always named after each question, but the order was varied from one respondent to another. It is recognized that Jews and Greeks may be native-born whites. If a respondent called attention to this fact, a rare occurrence, he was informed of the difficulty faced in finding exact criteria of group characteristics.

DISCUSSION

Table 1 presents a summary of the basic data obtained from the analysis of the questionnaire, from information given by parents about their occupation and education, and from a revision of the Chapin Social Status Scale (2).

Let us consider the social status of the family, the occupational level of the fathers, and the educational level of both parents before turning to the data directly bearing upon racial attitudes. It is immediately apparent from Table 1 that the white and Negro residents of District Y live in homes with better material equipment and having a higher esthetic level than the residents of District X, as judged by the two raters using the Chapin Social Status Scale. Although the white residents of District Y are rated higher than Negro residents, the difference in mean scores is not statistically significant. A similar comparison between white and Negro residents of District X shows a difference in mean scores which is significant at less than .01 level of confidence. The homes of Negroes in District X are rated as distinctly inferior to the homes of Negroes in District Y, for the difference in mean scores is significant at less than .01 confidence level. Although the white residents of District X generally have better homes than those of Negro neighbors, these white residents usually maintain dwelling places on a level inferior to those of white persons living in District Y, since the difference in mean scores has a confidence level of less than .01. Had no attempt been made to treat the two districts separately, it would have appeared that Negro residents did not differ significantly in social status from white residents, since the difference between the means of the combined scores of white parents and the combined scores of Negro parents does not meet a satisfactory level of confidence. We have established that the status of the home bears no clear relationship to racial attitudes.

TABLE 1

MEAN SCORES AND STANDARD DEVIATIONS ON QUESTIONNAIRE AND ON PERSONAL DATA FOR WHITE AND NEGRO PARENTS

VARIABLES INVESTIGATED	DISTRICT X						DISTRICT Y					
	MEN			WOMEN			MEN			WOMEN		
	N	MEAN	*SD*	*N*	MEAN	*SD*	*N*	MEAN	*SD*	*N*	MEAN	*SD*
Questionnaire Scored for Racial Attitudes												
Whites toward Negroes	68	18.15	6.19	78	17.53	5.60	60	18.00	6.51	66	19.44	6.53
Negroes toward Whites	3	28.00	3.56	6	29.17	1.39	16	30.37	0.86	18	29.22	1.79
Whites toward Jews	68	24.55	4.96	78	24.47	4.14	60	22.38	5.93	66	24.37	3.86
Negroes toward Jews	3	27.67	2.59	6	29.00	1.73	16	29.25	2.75	18	28.77	2.28
Social Status of Family (Chapin Social Status Scale)												
Whites	79	98.31	34.44				64	122.03	31.33			
Negroes	6	49.33	23.77				19	110.89	31.22			
Occupational Level of Father (Minnesota Scale of Occupations)												
Whites	68	3.98	1.36				60	4.20	1.50			
Negroes	3	5.33	0.77				16	5.06	0.91			
Educational Level of Father and Mother												
Whites	68	10.82	2.40	78	11.34	2.00	60	11.08	2.33	66	11.19	1.90
Negroes	3	9.33	1.72	6	9.50	2.36	16	11.87	2.52	17	12.00	1.97

Negro fathers are engaged in occupations usually regarded as requiring less skill and occupational intelligence than the occupations of the fathers of white children, according to ratings made on the basis of the Minnesota Occupation Scale (3). Because relatively few Negro fathers were interviewed, and in view of the similarity of the ratings of Negro men in the two districts, the ratings have been combined and compared with those of all white men in Districts X and Y. The difference in mean scores, significant at less than the .01 level of confidence, clearly portrays the difficulties Negroes face in job competition with white men. The white men in our sample usually held skilled jobs whereas the Negro men were more often employed in semiskilled occupations.

This discrepancy in occupational status cannot be accounted for in terms of formal education. In comparison with the 128 white men residing in the two districts, the 19 Negro men, as a group, have had slightly more formal education. The same trend is present when the 23 Negro women are compared with the 144 white women, for again the Negro women have had slightly more formal education than the white women. Both white and Negro residents of these two districts have had more formal education than the 8.9 years for persons 25 years and over in urban Minnesota.[8] The Negro residents of District Y do appear to represent a higher social status and a higher educational level than the Negro residents of District X, which is the less desirable residential area. In another article,[9] it will be shown that this kind of selection does not tend to reduce the antipathy of white neighbors.

Before we consider the data bearing upon racial attitudes, attention may usefully be called to a few descriptive categories helping to define the white families. In general, more white parents in District Y are newcomers than in District X: 23 per cent of the former and 13 per cent of the latter have resided in their neighborhoods less than two years. Yet a greater proportion of the parents of District Y owned their own homes than did the parents of District X, the percentages being 69 and 43, respectively. These differences in home ownership are consonant with the higher social status of the families in District Y. Of the parents who did not own their homes, 78 per cent in District X and 72 per cent in District Y said they would not consider buying homes in the neighborhood.

Parents' Attitudes

Satisfaction with the neighborhood was not as marked for the white parents of District X, the less desirable area, as for the white parents of District Y. All parents were asked if they liked or disliked living in the neighborhood. Only 51 per cent of men and women in District X, in comparison with 74 per cent of men and women in District Y, indicated that they liked living in the neighborhood. Dislike of the neighborhood, expressed by a parent, generally is indicative of more prejudice toward Negroes than is liking the neighborhood.[10] Yet only seven persons in District X and nine persons in District Y said they disliked their neighborhoods because Negroes had become residents in them.

Negro families own their homes and like the neighborhood to a greater extent than white families. Home ownership characterized 80 per cent of the Negro families and satisfaction with the neighborhood characterized 93 per cent of the Negro men and women. Needless to say, the appropriate comparisons with the white families yield differences in percentages that are significant at less than the .01 level of confidence. To live in these two neighborhoods, it appears that Negroes are usually required to purchase homes. The high level of satisfaction they display with their living conditions may be related to contrasts between them and those areas in the city where many other Negroes live, areas generally lower physically and occupied by persons with lower social and economic status.

Turning again to Table 1, under the caption *Questionnaire scored for racial attitudes,* we note that white men and women, as groups, do not differ markedly in their attitudes toward Negroes, nor are there any significant differences in the mean scores for the residents of both districts. It may be recalled

[8] *Statistical Abstract of the United States.* U. S. Department of Commerce, Bureau of Census, 1948.

[9] To be published as one of four articles dealing with aspects of racial prejudice. IV. Prejudice and discontent.

[10] *Loc. cit.*

that the range of attitude scores was from +7 to +31. High scores point in the direction of little or no prejudice. The mean scores of white parents typify persons who make voluntary comments to the effect that they fear marriage possibilities between their children and Negroes, or persons who say they would allow their children to play only when young with Negroes, or persons who say they will allow their children to play with Negroes only when in groups having a predominantly white membership.[11] Clearly, the white parents generally show more antipathy toward than regard for Negroes.

The questionnaires were also scored to obtain an index of prejudice toward Jews. In both districts, the antipathy of white men and women is greater toward Negroes than toward Jews, since each appropriate contrast yields a difference between mean scores which is significant at less than the .01 level of confidence. There were very few Jewish families among the respondents, hence these differences cannot be accounted for as a defense of the culture of a minority group. In only one instance do groups of white parents show disparity in attitudes toward Jews. The group of 60 men of District Y express the greatest degree of anti-Semitism. Comparisons between the mean scores of these 60 men and the mean scores of the remaining three groups of men and women yield differences which are significant at confidence levels of less than .01.

Negro residents differ markedly from the white parents in their attitudes toward other ethnic or racial groups. The Negro interviewers of Negro parents were particularly trained and instructed to secure answers as frank as possible and there is no reason to assume they did not do their best to meet requirements. Yet the extremely high scores of the Negro respondents, together with the very low variances, call into question the appropriateness of the questionnaire as an instrument for tapping prejudice toward white persons. This questionnaire was constructed to evoke responses from white adults and it was applied to Negro adults primarily to avoid accusations of discrimination rather than in the expectation that it would yield as valid results with Negro parents. The Negro interviewers and several other Negro citizens have said, "We have been inducted into tolerance through countless lectures and admonitions, so that we realize we must be tolerant to set an example to those who are less tolerant toward us. We have our own prejudices even though we do not think they are as great as those white people hold toward us." With this opinion we concur. It may be that the devising of a useful questionnaire for Negro parents will have to await the skill of a qualified Negro investigator, one who is sensitive, as white persons are not, to the intimate and subtle attitudes of Negro people in the northern part of the United States.

Table 2 assembles the responses of white parents that show a community of conventional stereotypes about Negroes. None of the questions appearing in this table was used to derive an index of prejudice.

There is surprisingly close agreement between the responses of white residents in Districts X and Y. The greatest discrepancy occurs in the answers to question 6, yet this difference seems a reasonable one in view of the greater proportion of parents in District Y who own their homes. The consensus of white parents is that Negroes are a threat to property values in a neighborhood having a preponderance of white residents.

Although a majority of the white respondents in both districts indicated that Negroes would have a deleterious effect upon home and property values, in District X 32 per cent and in District Y 45 per cent of them said that Negroes had already made property less valuable. Since between 34 and 38 per cent of the men and women interviewed did not recognize that Negroes lived in the neighborhood (see question 9), it seems reasonable to assume that in answering questions 4, 5, 6, and 9 the respondents were thinking of the neighborhood in a restrictive sense. Each district covered an area approximately 8×10 city blocks. Negro families tended to live in certain sections within each district even though their homes were interspersed among those of white families. It would be possible, therefore, for white parents to avow that Negroes did not live in their neighborhood, or to assert justifiably that Negroes

[11] *Op. cit.*, II. Methods for validating a questionnaire concerning White-Negro relationships by means of voluntary statements.

had not made property less valuable in the neighborhood.

Slightly more of the white respondents of District Y have invited Negroes into their homes, and have voluntarily called upon Negro neighbors than have the white parents of District X. These slight differences may be related to the higher educational and social status levels of the Negroes in District Y.

Children's Attitudes

Again let us consider first the supplementary and non-attitudinal measures. The Home Index (4) takes into account not only the physical features of the home but also the educational attainments of parents, together with their membership in clubs or organizations, such as study, art, or civic clubs. In general, the findings obtained by this index

TABLE 2

ADDITIONAL QUESTIONS ASKED OF WHITE RESPONDENTS IN DISTRICTS X AND Y BUT NOT USED IN COMPUTING A SCORE OF RACIAL PREJUDICE

QUESTION	REPLY	PERCENTAGE OF ALL RESPONDENTS	
		DISTRICT X	DISTRICT Y
1. Do you attend any church?	Yes	79	85
2. Are you a church member?	Yes	43	51
3. Which of these groups of people do you think most of your neighbors would object to?	Negroes	86	78
	Others *	2	6
	None	1	3
	Don't know	11	13
4. Which of these groups, if they lived in your neighborhood, do you think would make your home or property around here less valuable?	Negroes	94	95
	Others	1	2
	None	2	2
	Don't know	3	1
5. Which group do you think would hurt property values most?	Negroes	87	87
	Others	8	10
	None	2	2
	Don't know	3	1
6. Have any of these groups already made property less valuable in this neighborhood?	Negroes	32	45
	Others	1	0
	None	49	40
	Don't know	18	15
7. Have you ever invited members of the following groups into your home?	Negroes	19	25
	Others	49	35
	Whites only	31	40
	Non-committal	1	0
8. Have you voluntarily called upon, or dropped in to see and chat with any of your neighbors in the following groups?	Negroes	12	19
	Others	21	14
	Whites only	56	67
	Non-committal	11	0
9. Are there any families in these groups living in this neighborhood?	Negroes	66	60
	No Negroes	34	38
	Don't know	0	2

* Includes Japanese, Jews, Greeks, and Indians.

The main implication of the answers to questions 7 and 8 is that Negroes who live in a predominantly white neighborhood will not be accorded friendly and relatively free communication with the members of the majority group. Negroes may expect, however, to find white parents who have learned to judge persons apart from ethnic or racial group membership, parents who insist that they think in terms of individuals and who, therefore, ignore the pressures exerted upon them by white neighbors and friends.

parallel those presented in Table 1 under the headings *Social status of family* and *Educational level of father and mother*. The ratings obtained from the white children of District Y are higher than those obtained from the white children of District X, as shown in Table 3. The difference between the two means of 11.69 and 10.61 is significant at less than the .02 level of confidence. Negro children who live in the better residential area also make a higher rating on the Home Index than the Negro children who live in

District X, since the difference in mean scores is significant at a confidence level of less than .01.

In both districts, the white children have higher IQ's than the Negro children. The white children of District X have a lower level of general intelligence than the white children of District Y, for the difference in mean scores is significant at less than the .02 level of confidence. Likewise, the Negro children of District X achieve lower mean IQ's than the Negro children of District Y, the confidence level of the difference between the means falling below .05. These data support the interpretation made earlier, that the families in District Y tend to have higher social and educational status than the families in District X.

The behavior rating scales filled out by the teachers show that Negro children are considered to present more behavior problems than white children. Within each school, only the contrast between the white and Negro children of District X yields a difference in mean scores that meets a satisfactory level of statistical significance, namely less than .05. We did not find any significant relationships between these rating scales and the indices of prejudice.

Neither the attitude scale nor the questionnaire scored for racial attitudes toward Negroes yielded any important differences between the two groups of white children. In this respect, the children are like the adults of the two districts. Unlike the adults, however, the white children do not manifest more preference for Jews than for Negroes, since the mean scores on the questionnaire indicating attitude toward these two groups are almost identical. The white children have not learned to differentiate clearly, as have their elders, the two ethnic or cultural groups into a hierarchy of preference; if they had, the Negroes or the Jews would have been assigned consistently to a higher or lower position of choice. Because the schedules administered to adults and children posed different questions and situations, it is not possible to make direct comparisons between the mean scores of groups of children and groups of parents.

The Negro children hold more favorable attitudes toward white children than the latter do toward Negro children. It may be noted, however, that the Negro children with the lower mean IQ's and from the less desirable residential district hold more favorable attitudes toward white children than do the Negro children of District Y. The difference between the mean scores on the questionnaire

TABLE 3

MEAN SCORES AND STANDARD DEVIATIONS ON ATTITUDE SCALES AND QUESTIONNAIRE AND ON PERSONAL DATA FOR WHITE AND NEGRO CHILDREN

VARIABLE INVESTIGATED	DISTRICT X			DISTRICT Y		
	N	MEAN	*SD*	*N*	MEAN	*SD*
Attitude Scales Scored for Racial Attitudes						
White children toward Negroes	83	14.86	7.92	69	16.19	6.92
Negro children toward whites	10	24.10	1.97	21	21.57	3.37
Questionnaire Scored for Racial Attitudes						
White children toward Negroes	83	21.23	4.02	69	22.58	4.30
Negro children toward whites	11	28.27	1.59	22	25.59	3.18
White children toward Jews	83	21.81	3.70	69	21.87	3.12
Negro children toward Jews	11	21.18	2.48	22	21.68	3.88
Intelligence Quotient						
White children	81	105.48	10.84	63	109.79	10.81
Negro children	10	92.30	12.43	20	103.55	12.79
Behavior Rating Scale						
White children	82	68.39	17.69	69	74.61	23.53
Negro children	10	80.40	20.42	21	81.54	25.43
Home Index						
White children	83	10.61	3.14	69	11.69	2.56
Negro children	11	9.00	2.76	21	12.64	2.26

of these two groups of Negro children is significant at less than the .02 level of confidence. In attitudes toward Jews, Negro and white children are remarkably alike. The questionnaire scored for racial attitudes shows that Negro children manifest very little prejudice toward non-Jewish white children but considerable prejudice toward Jewish children. White children, as noted earlier, express prejudice toward Jews and Negroes to an equal and considerable degree.

Interrelationships of Racial Prejudice

Table 4 presents the correlations among the various measures of prejudice which may be considered statistically significant. Most of the correlations in the table are statistically significant at less than the .01 level of confidence. A few of these relationships seem to warrant discussion.

TABLE 4

COEFFICIENTS OF CORRELATION WHICH INDICATE STATISTICALLY SIGNIFICANT RELATIONSHIPS BETWEEN RACIAL PREJUDICE AND OTHER MEASURES

(White Respondents)

VARIABLE	*N*	*r*
Children's Questionnaire regarding Negroes and Child's IQ	144	.18*
Children's Attitude Scales regarding Negroes	152	.36
Questionnaire regarding Jews	152	.28
Occupational level of father	134	.21*
Father's Questionnaire regarding Negroes	134	.22*
Mother's Questionnaire regarding Negroes	151	.21
Children's Attitude Scales and Father's Questionnaire regarding Negroes	134	.27
Mother's Questionnaire regarding Negroes	151	.29
Children's Questionnaire regarding Jews and Last Grade Mother Completed	151	—.18*
Father's Questionnaire regarding Negroes and Questionnaire regarding Jews	128	.60
Boy's Questionnaire regarding Negroes	62	.36
Boy's Attitude Scales regarding Negroes	62	.36
Mother's Questionnaire regarding Negroes	127	.52
Mother's Questionnaire regarding Negroes and Questionnaire regarding Jews	144	.55
Girl's Questionnaire regarding Negroes	79	.24*
Girl's Attitude Scales regarding Negroes	79	.38
Father's Questionnaire regarding Jews and Mother's Questionnaire regarding Jews	127	.27

* Less than .05 level significance; all others less than .01.

The scores of the attitude scales, although they correlate positively with the scores on the children's questionnaire, appear to reflect different dimensions of prejudice from those tapped by the questionnaire. The questionnaire stressed what the white child does in his contacts with Negroes, whereas the attitude scale stressed what he thought about Negroes or what he would let Negroes do.

Parent-child resemblances in racial attitudes are not as close as other relationships involving attitudes toward institutions (6, p. 169). When the sex of the child is ignored, the children's scores on the questionnaire and the scores on the attitude scale, indicating orientation toward Negroes, correlate with the scores of mothers and fathers on the adult questionnaire between +.21 and +.29. The correlations between the scores on the questionnaire presented to fathers and the scores on the questionnaire and on the attitude scale obtained from sons are both +.36. Similar comparisons involving mothers and daughters yield correlations of +.24 and +.38, respectively. These coefficients are relatively low. In another article [12] emphasis has been put upon the disagreements existing between white parents and their children regarding the discouragement of communication of white with Negro children. It does appear that the white parents who were selected to represent the population of two school districts made little effort to inculcate racial tolerance in young children; contrariwise, in the majority of instances, these parents used neither strong nor consistent negative reinforcements when their children interacted with Negro children in school, on the playground, or in selected play groups. The ambivalent feelings of white parents toward Negro neighbors do not create a social atmosphere conducive to the establishment of a clearly delineated set of attitudes in young children.

Parents resemble each other more than they resemble their children in attitudes toward

[12] *Op. cit.*, III. The effect of parental discouragement of play activities upon the attitudes of white children toward Negroes.

Negroes. There were 128 fathers in the two school districts combined, but in one of these families the mother was dead. The correlation between the scores on the questionnaire of 127 husbands and wives was +.52. In contrast, these parents do not agree with each other in their attitudes toward Jews, +.27, as closely as they agree in their attitudes toward Negroes. Yet men and women tend to be prejudiced toward Jews if they are prejudiced toward Negroes. The questionnaire when scored for attitudes toward Negroes and Jews yielded correlations of +.60 for the fathers and +.55 for the mothers. That the attitudes of white children are not as definitely crystallized as those of their parents is evident from the correlation of +.28 between the scores on the questionnaire derived from the children's responses to Negroes and Jews.

A considerable number of variables, which on a priori grounds seemed to be related to racial prejudice, were found to have even less predictive value than those just reviewed. Certain of the coefficients assembled in Table 5 supplement interpretations already made on the basis of the previous set of statistically significant correlations. For example, the agreements in expressed attitudes toward Negroes of parents and children of the opposite sex are lower than the agreements of parents and children of the same sex. With respect to attitudes toward Jews, agreement between parents and children is less than the agreement existing when Negroes are the focus of consideration.

To complete the analysis of data, several other relationships were computed but only one of them reached a satisfactory level of statistical significance. The Home Index ratings yielded the following correlations: child's IQ, +.11; child's score on questionnaire regarding Negroes, −.06; child's attitude scale, −.08; and one coefficient of −.22, with a level of confidence of less than .02, with the occupational level of the father. In addition, it was found that the occupational level of the fathers correlated .00 with the IQ of the children. The fathers in the sample were relatively homogeneous with respect to occupational level, most of them being employed in skilled trades or jobs requiring approximately the same degree of intelligence. The correlation of −.22 between the Home Index, which was filled out by the children, and the occupational level of the fathers is not as unusual as it first appears. The men in District X, the physically less desirable area, held jobs which placed them slightly higher in the occupational intelligence scale than the men in District Y. The

TABLE 5

COEFFICIENTS OF CORRELATION WHICH INDICATE STATISTICALLY INSIGNIFACANT RELATIONSHIPS BETWEEN RACIAL PREJUDICE AND OTHER MEASURES

(White Respondents)

VARIABLE	*N*	*r*
Children's Questionnaire Regarding Negroes and		
Haggerty-Olson-Wickman Rating Scale part B–3	151	.04
Haggerty-Olson-Wickman Rating Scale part B–4	151	.00
Last grade father completed	134	.04
Last grade mother completed	151	.14
Children's Attitude Scales and		
Haggerty-Olson-Wickman Rating Scale part B–3	151	.01
Haggerty-Olson-Wickman Rating Scale part B–4	151	−.04
Occupational level of father	134	.02
Last grade father completed	134	−.01
Last grade mother completed	151	.01
Children's Questionnaire Regarding Jews and		
Haggerty-Olson-Wickman Rating Scale part B–3	151	−.05
Haggerty-Olson-Wickman Rating Scale part B–4	151	−.10
Occupational level of father	134	.02
Last grade father completed	134	−.04
Father's Questionnaire regarding Jews	128	.14
Mother's Questionnaire regarding Jews	144	.15
Father's Questionnaire Regarding Negroes and		
Occupational level of father	134	.01
Last grade father completed	134	.02
Girl's Questionnaire regarding Negroes	72	.05
Girl's Attitude Scales regarding Negroes	72	.16
Mother's Questionnaire Regarding Negroes and		
Boy's Questionnaire regarding Negroes	72	.16
Boy's Attitude Scales regarding Negroes	72	.19
Last grade mother completed	151	.13
Father's Questionnaire Regarding Jews and		
Boy's Questionnaire regarding Jews	62	.19
Girl's Questionnaire regarding Jews	72	.08
Mother's Questionnaire Regarding Jews and		
Boy's Questionnaire regarding Jews	72	.17
Girl's Questionnaire regarding Jews	79	.09

men in District X lived in homes rated by interviewers as less desirable, or indicative of a lower social status, than those occupied by men of District Y. Home ownership also was not as characteristic of the men of District X as of the men of District Y. Even though the predictive value of the correlation is negligible, its direction is defensible on logical grounds.

SUMMARY AND CONCLUSIONS

In neighborhoods undergoing an infiltration of Negro citizens, white residents express markedly different racial attitudes. A small proportion of the white adults, approximately 20 per cent, showed few or no antipathies toward Negroes in answering a questionnaire. A larger proportion were definitely antagonistic and resentful of what they considered infringements by Negroes upon the rights of white people. These persons often seized opportunities when answering questions to unburden themselves of their racial antagonisms. The majority of the respondents, however, were ambivalent in their attitudes, sometimes regretful because they felt unable to overcome what they recognized as "unfounded prejudices," and sometimes chagrined because the proximity of Negroes had made them aware of hitherto unconsciously expressed prejudices. Under the varied circumstances characterizing the social atmosphere of the homes of these adults, many children are not afforded a consistent and clearly evident set of evaluations of Negroes, with one result being that the white children do not closely resemble their parents in attitudes toward Negroes even though the parents share attitudes having considerable similarity.

Although the inquiry was fashioned to underemphasize interest in Negroes by mentioning them in conjunction with other racial and ethnic groups, the predisposition of the majority of white respondents led them to put the Negroes into the focus of attention. In view of this inclination, the fairly high correlations established between attitudes toward Negroes and Jews take on added meaning. These correlations support the contention that adults who are prejudiced toward Negroes are most likely those who will be prejudiced toward Jews. Perhaps the opposite form of statement should also be made, for too often we center attention upon the neglect of social responsibility rather than upon the observance of it. The white adults who had learned to think in terms of the attributes of individuals, instead of the imputed and mythical attributes of groups, could be said to hold more favorable than unfavorable attitudes toward both Negroes and Jews. In contrast, the white children were not as neatly stylized in their attitudes toward Negroes and Jews, since, as a group, they had not developed a clear order of preference between them.

A number of variables were shown not to bear any clear-cut relationship to racial prejudice as expressed by white adults, among them being social status, educational level, occupational status, and affiliation with a church or attendance at church services. At least two well-established institutions, religion and education, seem to have very little impact upon the fashioning of a consistent orientation in white persons living in a northern metropolis toward accepting Negro and Jewish citizens in terms of individual worth.

Even though the questionnaire did not appear to be entirely appropriate for measuring the attitudes of Negro adults toward white persons, it did provide important data indicating that they had achieved a level of formal education approximately equal to that of the white respondents, and that, in one district, their level of social status, as judged by the arrangements and cultural objects within the homes, was not significantly different from the social status of white persons. It seems, however, that Negro men have not been given adequate opportunities to utilize their education in appropriate employment. Most of the Negro men were employed in positions requiring less skill than they seemed capable of manifesting.

The Negro children held more favorable attitudes toward white non-Jewish children than they did toward Jewish children. Both the white and the Negro children were remarkably alike in their expression of attitudes toward Jewish children. Since very few Jewish children lived in either of the two districts, it appears probable that both white and Negro children were oriented in their responses by an unconsciously held but subtle

and pervasive stereotype of what Jewish children are expected to be like.

It has not been possible to discuss all aspects of this inquiry into racial prejudice. In succeeding accounts, other facets of prejudice will be presented. In the light of data issuing from this inquiry and of the data already made available by other investigators, it seems as if we have barely come to grips with the problems of racial attitudes. Before we can attempt to modify attitudes with hopes of success, it will be necessary to plumb the dimensions of prejudice by more exacting and penetrating methods than have been used. Such an inquiry as we have undertaken probably has most value in selecting persons for intensive interviewing and testing because they typify many other individuals along the continuum of social intelligence, which ranges from a rare or a unique form of discrimination to a common or simple form of generalization.

REFERENCES

1. BETTELHEIM, B., & JANOWITZ, M. *Dynamics of prejudice.* New York: Harper, 1950.
2. CHAPIN, F. S. *The measurement of social status.* Minneapolis: Univer. of Minnesota Press, 1933.
3. GOODENOUGH, FLORENCE L., & ANDERSON, J. E. *Experimental child study.* New York: Century, 1931. Pp. 234–238, 501–512.
4. GOUGH, H. G. A short social status inventory. *J. educ. Psychol.*, 1949, **40**, 52–56.
5. GOUGH, H. G., HARRIS, D. B., MARTIN, W. E., & EDWARDS, M. Children's ethnic attitudes: I. Relationship to certain personality factors. *Child. Develpm.*, 1950, **21**, 83–91.
6. HARRIS, D. B., GOUGH, H. G., & MARTIN, W. E. Children's ethnic attitudes: II. Relationship to parental beliefs concerning child training. *Child Develpm.*, 1950, **21**, 169–181.
7. HARTLEY, E. L., ROSENBAUM, M., & SCHWARTZ, S. Children's use of ethnic frames of reference: An exploratory study of children's conceptualizations of multiple ethnic group membership. *J. Psychol.*, 1948, **26**, 367–386.
8. LASKER, B. *Race attitudes in children.* New York: Holt, 1929.
9. LIPPITT, R., & RADKE, M. New trends in the investigation of prejudice. *Ann. Amer. Acad. Polit. soc. Sci.*, 1946, **244**, 167–176.
10. RADKE, M. J., & TRAGER, H. G. Children's perceptions of the social roles of Negroes and whites. *J. Psychol.*, 1950, **29**, 3–33.
11. RADKE, M. J., TRAGER, H. G., & DAVIS, H. Social perceptions and attitudes of children. *Genet. Psychol. Monogr.*, 1949, **40**, 327–447.

Received March 1, 1952. Prior publication.

THE VALIDITY OF HYPNOTICALLY INDUCED COLOR HALLUCINATIONS

BERNARD G. ROSENTHAL AND HOWARD MELE [1]

University of Chicago

When the hypnotist suggests to his hypnotized subject that there is a colored card before him, why does the good subject agree to the presence of the nonexistent color? Does he affirm its presence because he really sees it or does he merely give verbal agreement without actually experiencing its existence?

This paper is concerned with the validity of hypnotic subjects' reports of hallucinated colors. Do they really see them or are they merely saying that they do? Previous attempts to answer these questions have been reported with contradictory results by Erickson and Erickson (3), Hibler (4), and Binet and Féré (1). Unfortunately, however, the methods and design utilized in many of these studies have such serious defects that their conclusions are in doubt. This will be discussed later in detail.

Procedure

In the hope of remedying these defects, the method used in this study for answering the questions posed above consisted in testing for the occurrence of after-images on gray cards of colors previously hallucinated on gray cards. For example, when the *S* says he sees red in response to the hypnotist's suggestion, then this method of validating the *S*'s report would require him to see the negative after-image of red on a gray background immediately after he reported seeing red. If he reports this negative after-image correctly and if he is naive as regards verbal conscious knowledge of after-images, then there would be reason to believe that he did really see the color the hypnotist suggested and was not merely agreeing verbally. An alternative explanation of this same phenomenon involving unconscious verbal associations will be discussed later. In line with the assumptions underlying this method, the experimental procedure was as follows:

1. The *S*s were selected on the basis of being able to reach a deep hypnotic trance and of having visual hallucinations, including those of hypnotic, post-hypnotic, negative, and positive nature.

2. Each *S* was put into a deep state of hypnosis, involving at least 30 minutes of hypnotic suggestion.

3. Then the following statement was made:

> I'm going to show you a blue card. I'm going to hold the blue card in front of you. I want you to stare at the black dot in the center of the card for one minute. After the minute is up, I will take the blue card away and hold up an entirely new card. Then stare at the dot in the center of the second card, and tell me what you see. Understand? O.K. Now stare at the blue card.

A gray card 6 in. by 9 in. was now presented 24 in. from *S*'s eyes and held there for one minute. After the first ten seconds of its presentation to *S*, the following statement was made: "Do you see the blue card clearly? Keep staring at the dot in the center." After the minute was up another card was presented, identical in color with the previous one, but 9 in. by 12 in. in size. The second card was presented for two minutes at the same distance of 24 in. If *S* did not offer any comment after looking at the second card for 30 seconds, he was asked, "What do you see? Do you see anything?"

This procedure was repeated for red, yellow, green, and orange. Each time *S* was told by the hypnotist what color the first gray card was supposed to be.

There was a rest period of three minutes between each of the five pairs of presentations.

4. The subjects were awakened with amnesia suggested for the hypnotic state.

5. The whole experimental procedure as outlined above was repeated in the waking state as a control on the reported after-images in the hypnotic state.

6. The subjects were then asked in the waking state:

> If you stared at a green card for one minute, and the card was suddenly removed and you looked at a grey card, what would happen? Would you see anything? Would you see a color? What color?

This was repeated for the colors yellow, red, blue, and orange. This was followed by further questioning to determine if there had been any previous experience with after-images. The whole object of this questioning was to determine whether *S*s were naive as regards conscious knowledge of after-images so that the results in the hypnotic state might not be attributed to such knowledge. The questioning proceeded in such a way that suggestions as to what response might occur (even a colored one) as a result of looking at the second white card were avoided till the very last.

7. The subjects were rehypnotized and each of the real colors was presented, each followed by a gray

[1] The senior author is responsible for the conception and design of the experiment and the form of its report and interpretation presented herein. The junior author carried through the hypnosis of the *S*s and the experimental details associated with it (all under the supervision of the senior author).

card. The procedure was exactly the same as in the part reported under hypnotic suggestion of colors except now the colors were real. The rationale for this part of the study was to compare under hypnosis the after-images to true colors with hallucinated ones.

8. With amnesia suggested for the hypnotic state, the *Ss* in the wakened state were presented with each of the real colors followed by the gray card. The procedure was the same as in Part 7, except it was done in the waking state. The rationale for this part was to determine what the after-images to the true colors were in the waking state as contrasted to after-images to both true and hallucinated colors in the hypnotic state.

RESULTS

The results indicate (Table 1) that of the four *Ss*, three experienced negative after-images for every color hallucinated under the hypnotic trance while the fourth experienced negative after-images for four of the five colors hallucinated. Moreover, the negative images to the hallucinated colors were the same as the *Ss* experienced in the waking and hypnotic state when the real colors were presented with the exception of one color for the fourth subject, whose case is reported in detail below. According to the *Ss*' reports, the images to the real and hallucinated colors were the same in both hue and other qualities with the sole deviation, again, of the same single color for the fourth *S*. All attempts to obtain hallucinations in the waking state failed. The after-images of the real colors presented in hypnosis and in the waking state were of the normal color except for the same color for the fourth *S*.

The fourth *S*, who got all the negative after-images except one during hypnotic hallucination, failed on the after-image to blue. He described the card presented after the blue hallucination as a "very faint bluish-gray." When real blue was later shown to him, he said the after-image was gray. When questioned further, he said there might be a small amount of yellow in it. Moreover, blue objects seem much darker to him than to others. His after-images to yellow were described as blue. He was completely naive when tested for knowledge of after-images.

One *S* had knowledge of after-images. Two after-images to hallucinated colors, however, were different from the answers he gave on the verbal test. He said he would expect the after-image of green to be pinkish but for the after-image of this hallucinated color he gave red. Similarly he said he would expect the after-image of orange to be blue, but he actually reported the after-image to be blue-green for both hallucinated and real orange. Thus it appears that even when an *S* has verbal knowledge of after-images this does not influence his reported after-images to real colors or to hypnotically hallucinated ones. Hence, of the 20 possible correct hallucinated after-images for all *Ss*, only one was incorrect, and that for an *S* who usually in the waking state was defective for vision of the color hallucinated and also defective for the after-image of this color under normal circumstances. Again, even in the case of the *S* who had knowledge of after-images,

TABLE 1

RESULTS OF EXPERIMENTAL AND CONTROL SERIES

Subjects 1–4 *

COLOR STIMULUS	EXPERIMENTAL SERIES RESPONSES TO COLOR AS VERBAL STIMULUS (HALLUCINATED COLOR) (Hypnotic state)	CONTROL SERIES I RESPONSES TO COLOR AS VERBAL STIMULUS (HALLUCINATED COLOR) (Waking state)	CONTROL SERIES II AND III RESPONSES TO ACTUAL COLOR STIMULUS (Waking and Hypnotic states)
Red	Green	Gray	Green
Blue	Yellow *(Bluish-gray for *S*–4 only)	Gray	Yellow *(Grayish-Yellow for *S*–4 only)
Yellow	Blue	Gray	Blue
Green	Red	Gray	Red
Orange	Blue-Green	Gray	Blue-Green

though his knowledge was incorrect in two cases, his actual after-images to hallucinated colors were correct. This appears to mean that verbal conscious knowledge of after-images does not influence responses to after-images of hypnotically suggested colors.

DISCUSSION

The above results are in line with the findings of Binet and Féré (1) and Erickson and Erickson (3), though each of these studies differs from ours in many significant respects. Our results are not compatible, however, with those of Hibler (4). Though Hibler's methods differed from ours in some respects, it would pay to consider briefly certain crucial points in his study (as well as relevant criticisms by Erickson (2) of Hibler's research) not only to account for the differences between his data and ours but, more important, because they are among the serious dangers that must be carefully avoided in a field where the slightest underestimation of the sensitivity and suggestibility of the *S*s to even the most subtle cues leads to disastrous results.

Those points in Hibler's study that are so important and that may account for the differences in results are:

1. Briefness of time spent in inducing the trance. Five minutes is much too short a time to expect significant transformations from waking state to deep hallucinated trance. If his *S*s gave only verbal assent to the hypnotist's suggestions of hallucinated color, as Hibler is inclined to believe, this may be true only because the basic changes in the hypnotic *S*, indispensable for a valid test of hallucinatory experience, were not given sufficient time to occur in Hibler's study.

2. Hibler's directions to his *S*s and his other experimental contrivances set up an experimental atmosphere that effectively nullified an adequate demonstration of the phenomenon to be studied. Among the more salient of these misdirected techniques were:

a. The test to determine *S*'s knowledge of after-images was given before the actual experiment and hence may have oversensitized him to the phenomenon of after-images, thus causing him to have expectations of these phenomena in the subsequent hypnotic experiments. These expectations, in turn, may have influenced his responses in the hypnotic hallucinatory series in the direction of his pre-experimental after-image test answers and hence contaminated the spontaneous responses in the critical hypnotic phase of the study.

b. By employing deceptive cues, requesting certain responses and then inhibiting them, and utilizing contradictory suggestions to his *S*s, Hibler did not really get an "acting upon his suggestions" as Erickson points out, "but only an abidance by them. It is not sufficient to have the subject agree only—rather the subject must perform the infinitely harder task of experiencing" (2, p. 168). By using these imprecise techniques it may be argued that Hibler might have failed to elicit such "experiential" responses in his *S*s and hence failed to elicit behavior that was relevant to the main purpose of the experiment.

Though Erickson and Erickson's (3) study (which presumably confirmed the validity of hypnotic hallucinatory experience) is sophisticated in the subtleties of hypnotic technique and indicates awareness of how unintentional subtle deception, contradictory suggestions, etc., may misdirect an adequate investigation of hypnotic phenomenon, their own experimental design is defective in exactly those respects where Hibler's is not, i.e., precise experimental control and methodology—and these errors may tend to throw some doubt on the results obtained. Among the more salient of these limitations are:

1. A test to determine knowledge of after-images was given before the actual experimental series. As in the case of Hibler's study, this may have forewarned the *S*, consciously or unconsciously, of what was to come, reactivated old verbal associations and other traces between verbal color stimulus and responses, and in other respects influenced the *S*'s spontaneous responses when the critical experimental tests were made under hypnosis.

2. Related to this is the fact that Erickson and Erickson gave directions to their *S*s specifying that the second sheet they would see in the hypnotic experimental series would be a color. There is at least room for suspicion that this device may have given the *S*s definite expectations that their responses must be colored ones (and not black or white); and this expectation taken in con-

junction with the responses on the after-image test might have modified the spontaneous responses made during the hypnotic experimental series; at the very least, it might have facilitated the eliciting of unconscious verbal associations to the verbal color stimulus presented during hypnosis, and thus have had much the same effect.

3. Erickson and Erickson do not report the duration of time the *S* was required to fixate the first presented sheet. Was it constant or variable from *S* to *S*? If constant, how long was it? Clearly this variable may have affected the results obtained in several ways.

4. The word association test that the Ericksons used to determine knowledge of after-images was, as Hibler (5) points out, unstandardized, and, in general, a not exact method for stating confidently that the *S*s were naive as regards this knowledge.

5. Finally, there is no indication what the responses of the Ericksons' *S*s would be to the actual colors in the waking and hypnotic states. Lacking this important control, the results of their experiment might be open to considerable question.

The upshot of this evaluation of Hibler's and Erickson and Erickson's study is to point up how, in this difficult area of research, one must be particularly on guard against looseness of experimental design and against subtle nuances in directions to the hypnotic *S* which may influence results immeasurably by unwittingly utilizing this extreme suggestibility to reformulate and hence misdirect the hypnotic situation from a valid investigation of the hypnotic problem that is being considered.

If the study reported here has been reasonably successful in avoiding at least the most significant pitfalls of research in this area, the question arises as to how the results may be explained. One possibility is that the phenomenon reported here is essentially a central one, i.e., that under hypnotic auditory suggestion central visual processes are stimulated in some fashion and color is experienced thereby, although no externally adequate stimulus is present. When this visual experience, centrally originated, is projected on a gray background, the result is the negative after-image. The central origin of the color would be similar to the origin of colored dreams or similar color phenomena without adequate external stimulation. Whether or not such centrally originated processes would stimulate those peripheral processes which actually give rise to color, or whether the whole process is exclusively central is a question that clearly goes beyond our data. No more may be said here except that, according to this explanation, central processes via auditory stimulation have a crucial part in starting the whole phenomenon.

Another possible explanation of these phenomena bases them on unconscious verbal association. Although the *S* may have had many experiences with after-images, he may be completely unaware of them in actual experience or have no knowledge of them verbally. Yet connections are established between seeing a color, its after-image, and the verbal stimuli, i.e., red or green, that are associated with the color and its after-image. These connections, however, are below the threshold of recall for normal waking state. When hypnotized, the *S*, without experiencing the color, has his threshold for such associations sufficiently lowered so that he may remember what the proper connections are between the color suggested and its after-image, which he has already experienced innumerable times. The function of hypnosis is merely to lower the threshold for memory of these already established verbal associations and make evident that the *S* knows what he denies knowing in the waking state. Although all *S*s except one were naive as to knowledge of after-images, this does not mean that under hypnosis they may not have recalled this knowledge. That this *S*'s knowledge of after-images (i.e., the only *S* who knew about them) did not conform with his actual after-images either to real or hallucinated colors may only mean, according to this explanation, that the unconscious verbal connections he established were in line with his actual experience and hence more potent than his defective conscious verbal knowledge, which had little influence in determining recall in the trance because of its contradiction with the *S*'s actual experience. Since our *S*s were quizzed in the waking state and not in the hypnotic state, it might be possible that they were not as naive about knowledge of after-images, i.e., uncon-

sciously, as the test of conscious knowledge would indicate. At any rate, questioning in the hypnotic state about knowledge of after-images clearly represents a necessary step in the future of this kind of work. Ultimate disposal of this explanation, therefore, must be delayed until the just previously suggested research step is performed. The difficulty with this step is that one must be careful not to quiz the *S* about after-images before the colors are suggested, since that may stimulate him to expectations of certain kinds of reaction about after-images that would destroy the spontaneity of his responses. On the other hand, if one quizzes him afterwards in hypnosis about these phenomena, there is always the danger that under hypnosis he will recall the experiences he has just had with the hallucinated colors since while amnesia may be had in the waking state for hypnotic experience, it is extremely difficult to obtain amnesia in the hypnotic state for hypnotic experience.

Other possible approaches to the problem are: (*a*) employing more subtle and controlled deception of *S*s, (*b*) work with young children, where the connections between verbal stimuli and after-images are presumably not as well established as in adults, (*c*) various other experimental methods and designs, including those controlling the effects of pre-experimental series after-image tests on the subsequent hypnotic hallucinatory trials. Whatever methods may be ultimately employed, however, one can see the difficulties and dangers in this work where seemingly meaningless words and suggestions or supposedly rigorously controlled techniques take on the greatest unpredictable significance to the *S* and give results that really do not test the fundamental hypotheses the experimenter may have been interested in.

REFERENCES

1. BINET, A., & FÉRÉ, C. *Animal magnetism.* London: Kegan Paul, Trench, Trubner, 1891.
2. ERICKSON, ELIZABETH M. Critical comments on Hibler's presentation of his work on negative after-images of hypnotically induced hallucinated colors. *J. exp. Psychol.*, 1941, **29**, 164–170.
3. ERICKSON, M. H., & ERICKSON, ELIZABETH M. The hypnotic induction of hallucinatory color vision followed by pseudo-negative after-images. *J. exp. Psychol.*, 1938, **22**, 581–588.
4. HIBLER, F. W. An experimental investigation of negative after-images of hallucinated colors in hypnosis. *J. exp. Psychol.*, 1940, **27**, 45–67.
5. HIBLER, F. W. Note on Mrs. Erickson's comments on Hibler's work on negative after-images of hypnotically induced hallucinated colors. *J. exp. Psychol.*, 1941, **29**, 170–172.

Received April 7, 1952. Prior publication.

THE HONI PHENOMENON: A CASE OF SELECTIVE PERCEPTUAL DISTORTION[1]

WARREN J. WITTREICH
Princeton University

STABILITY and continuity are the rule rather than the exception in ordinary perception. In spite of constantly shifting impingements upon its sense organs, the human organism does perceive the world in an orderly and sensible manner. However the facts of everyday perception are not explanations, and it is often only by deliberately distorting or changing the commonplace that we can shed any light on the processes which allow anything to become "commonplace" in the first instance. Perceptual distortions have therefore come to play an important role in the understanding of basic perceptual processes. Recently a number of devices have been developed which deliberately place cues in conflict to such an extent that strikingly dramatic perceptual distortions can be obtained.

A number of distorted rooms have been constructed in which the floor slopes up to the right of the observer, the rear wall recedes from right to left, and the windows are of different sizes and trapezoidal in shape. When an observer looks at any one of these rooms from a certain point with one eye the room appears normal or almost normal, as if the floor were level, the rear wall at right angles to the line of sight and the windows rectangular and of the same size.[2]

One of these rooms has been built large enough so that people can enter the room and walk about in it. What the observer typically sees when he watches someone moving about in this room is a striking alteration in the observed individual's size. When standing in the corner to the observer's left, the person appears abnormally small; in the other corner he appears abnormally large. When walking from one corner to the other he appears to grow or shrink in size, depending upon the direction in which he is travelling. A smaller model of this room permits hands or faces to appear through the rear windows. As with the larger room, the hands or faces appear abnormally large or abnormally small, depending upon whether they are in the window to the observer's right or left.

The study to be reported here received its impetus from an unusual instance in which this typical pattern of observation was not reported. In 1949 a woman observed the faces of her husband and another man through the rear windows of the smaller room. The face of the other man was described as distorted in the usual manner, but no size changes whatsoever were reported for the husband; his face was described as being perfectly normal no matter which window it appeared in. Similar results were obtained in the large room. Again the other man was described in the usual manner: he appeared to grow or shrink and looked large or small, depending upon the corner in which he was observed. But again no such size changes were reported for the husband. No matter which corner he stood in, he was reported as looking perfectly normal—his usual size. This unusual observation was named the "Honi phenomenon" following the family nickname for the woman who first experienced and reported it.

This instance of a deviation from the typical pattern of observation of people in the room raises at least two specific questions. Is this observation simply a unique and isolated case or is it reproducible with other individuals? What are the psychological conditions which account for such a performance?

This paper reports two experiments which answer the first question by producing the Honi phenomenon in unmistakable fashion

[1] The work reported in this paper has been done under contract Nonr-27014 with Princeton University, a project initiated by the Professional Division, Bureau of Medicine and Surgery. The opinions expressed are those of the individual author and no not represent the opinions or policy of the Naval Service.

[2] These rooms were designed by Adelbert Ames, Jr., as part of a series of perceptual demonstrations. This description is modified from ITTELSON, W. H., & KILPATRICK, F. P. Experiments in perception. *Scientific American*. August, 1951, **185**, 50–55. See also CANTRIL, H. *The why of man's experience*. New York: Macmillan, 1950.

with a number of other people. The obvious and outstanding relationship between the woman and her husband in the original situation was simply the fact that they were married. It should also be noted that both were over sixty years of age, and that he was a very distinguished man whom she greatly admired and to whom she was devoted. The assumption was made that if the phenomenon has been observed in this couple, it might very well be observed in certain other couples. A group of couples, the majority of whom had been married for a relatively short time, were selected as subjects primarily because of their availability and willingness to participate in the experiment.

In the first experiment an attempt was made to reproduce conditions under which the phenomenon had originally been observed, with the addition of needed experimental controls. All that was required of the *S* was a description of people observed in the room. It was assumed that if a description of what appeared to be a case of selective perceptual distortion had been obtained in the in the original situation, it could be obtained again. The prediction was simply that an individual, when observing both a stranger and his or her marital partner in the room, would report less distortion in the description of the marital partner than in the description of the stranger. The second experiment was an attempt to confirm what was found in the first experiment, and to obtain a quantitative measure of the difference in the relative distortion of the marital partner as compared to the stranger.

EXPERIMENT I [3]

Ten married couples provided a sample of 20 *S*s in this experiment. At the time of the experiment six of the couples had been married less than one year. The remaining four couples had been married two, three and a half, five, and ten years, respectively. (See Table 1.)

Both the small and the large rooms were employed in this experiment. Both rooms were viewed with one eye only. With the small room, the *S* was asked to describe: (*a*) the room itself, (*b*) the hands of the experimenter placed through the rear windows, (*c*) a marble which ran across the room and gave the appearance of rolling uphill, (*d*) two situations with two people putting their heads through the rear windows: two strangers, and one stranger and the marital partner. (Each person was seen once in each window in each of the above situations.) Descriptions *a*, *b*, and *c* were requested primarily to see if the *S* was initially observing the room and objects in the room in the typical manner.

With the large room the *S* was asked to describe: (*a*) the room itself; (*b*) two situations with two people standing in the corners of the room: two strangers, and one stranger and the marital partner (each person was seen once in each corner in each of the above situations); and (*c*) two different people walking from the corner on the observer's left to the corner on his right and back again: a stranger, and the marital partner. The entire experiment was recorded on a wire recorder and transcribed.

RESULTS

An analysis was then made of each individual protocol to determine whether or not the *S* had spontaneously described a definite difference between the appearance of the marital partner and the appearances of both of the strangers, and whether this difference was in the predicted direction. Table 1 indicates that 6 out of the 20 *S*s did report a difference in the expected direction in the small room and 7 out of 20 reported a similar difference in the large room. Not once did an *S* report a difference other than in the expected direction, i.e., not once was the marital partner described as being more distorted than either or both of the strangers. It is also of interest to note that, although the exact same people did not display the phenomenon in both rooms, there are a few striking cases of overlapping. Taking both situations together, it can be seen that at least one member of all the couples married under one year displayed the phenomenon to some degree in one of the two rooms, but, with a single exception, no member of the couples married over one year displayed the phenomenon. That single exception was a man married only two years.

Because of space limitations only selected portions of the individual protocols which contain material specifically relevant to the hypothesis will be presented. In reading these protocols it is worth noting a number of things. First of all, when an *S* reports a difference, he does so in a spontaneous manner, and he, himself, often appears to be struck

[3] The experiment reported here is part of a larger pilot study which is reported in detail in KILPATRICK, F. P. (Ed.) *Human behavior from the transactional point of view*. Hanover, New Hampshire: Institute for Associated Research, 1952.

by both the existence of this difference and the direction it takes. Second, it will be noted especially in protocols D and E in the large room that while the marital partner is maintaining some degree of constancy, the room itself is undergoing a process of increasing distortion. It appears almost as though the *S* is faced with a choice of distorting either the marital partner or the room, and he chooses to distort the latter. In all of the following protocols the marital partner is capitalized.

Small Distorted Room

A. Male, married 3 months: HER'S looks bigger, but it doesn't look as much bigger than Bruce's did to your's although the windows look the same. It still looks bigger, but it doesn't look Bruce's head looks larger but NANCY'S looks—your (NANCY) head doesn't seem to have shrunk much. Bruce's head has grown.

E. Female, married 11 months: DAVID's head looks very big now. And Bruce's head looks the way it did before—when I looked at it the first time. Your head looks, darling; darling, you can't even get your head through that thing. But uh, I guess that's all. They both look in proportion. (Head's reversed) Gee. Golly. Well Bruce's head looks bigger now than it did the last time when he had his head on this side, and uh, and DAVID, I don't know about you. Your head looks big too, but you look normal in size—I think—whereas Bruce looks bigger than he is. I think that's all.

F. Male, married 11 months: Well, now it's reversed. As a matter of fact, Bruce looks larger in comparison to CHARLOTTE than CHARLOTTE did to Bruce in the previous situation. Other than size I can't see anything unusual. The faces appear

TABLE 1

MEMBERS OF MARRIED COUPLES REPORTING DIFFERENCES BETWEEN MARITAL PARTNER AND STRANGER WHEN OBSERVED IN THE DISTORTED ROOMS

COUPLE	LENGTH OF MARRIAGE	DIFFERENCE IN APPEARANCE OF MARITAL PARTNER AND STRANGER REPORTED BY HUSBAND OR WIFE *			
		LARGE ROOM		SMALL ROOM	
		HUSBAND	WIFE	HUSBAND	WIFE
1	3 mo.	x		x	
2	4 mo.	x		x	
3	7 mo.		x		x
4	3 mo.	x			
5	11 mo.	x	x	x	x
6	11 mo.			x	
7	2 yr. and 3 mo.	x			
8	3 yr. and 6 mo.				
9	5 yr.				
10	10 yr.				

* All differences reported are in the predicted direction, i.e., marital partner distorting less than stranger.

a great deal bigger. I think her head is bigger than Bruce's anyway naturally.

B. Female, married 7 months: I can't get over it. BOB'S head appears to be far away and Bruce's head appears to be close, and BOB'S head appears normal—uh, fairly normal—it's small, but it's more in proportion than Bruce's. That's all.

C. Male, married 4 months. BUNNY'S head looks normal, but Bruce's head is much larger than you would expect. (Heads reversed) Now BUNNY'S head is larger too, because of the size of the window; but not so much as it was the other way around. Bruce's head looked much larger than BUNNY'S does.

D. Male, married 11 months: I guess your (NANCY) head has grown, but it hasn't grown much. You're (Bruce) like you were before, I guess. I guess NANCY is mostly bigger and Bruce looks smaller and NANCY'S head is closer to me. (Heads reversed) That's right. That's amazing. normal, except Bruce seems closer and his face looks larger.

Large Distorted Room

A. Male, married 4 months: BUNNY'S head looks about the size or smaller than this light bulb in the room, and Bruce's looks twice as large. Bruce looks like he's about—yeah, yeah he looks twice as tall as BUNNY. Twice as large, not only in height, but in width and everything. And uh—the left window's grown larger than the right. Bruce looks much—of course that is in relation to the people—the left window dwarfs BUNNY and the right window is dwarfed by Bruce. When I look at the feet, the slant isn't as pronounced as you might think it is. But it is a terrific slant. But it doesn't look too bad from here. (Change places) Oh geez. It's the same relationship, but BUNNY looks—Bruce looked much taller than he should have—and BUNNY looks about her size. I mean I could get in and stand beside BUNNY

and look the same way she does—about the same size. Bruce looked much taller than BUNNY does, but Bruce at the same time looks as small as BUNNY did. The same holds with his head in relation to the light bulb and BUNNY'S too. But the fact is, I think more of it is tied up with the fact that you know one person so much more, you know. I mean I can put myself in BUNNY'S place and I know darn well that—uh—that's just about the size she would look. Doesn't make her look much larger, but Bruce it makes look smaller. Bruce looked like a giant standing in the corner and BUNNY looks normal, but at the same time I know darn well Bruce can't be that much smaller than BUNNY is. I mean on the street for example. She doesn't look like a giant and Bruce did. Put it that way. And now BUNNY looked like a midget and so does Bruce. Geez when I think back—even smaller he looks. I have to look at BUNNY in that corner again to tell you if Bruce looks smaller.

(Walking from corner to corner—marital partner) Well, BUNNY went from—everybody looks the same in that other corner—but BUNNY went from half her size to her normal size. I mean she looks normal now; she doesn't look any larger than she usually does. She got larger, but she didn't reach proportions which were beyond her normal. Geez, that's amazing. She doesn't start to get very much smaller until about two steps away from that very far corner there. All of a sudden she gets real small. And she's much further away now, too. Everybody looks the same size in that corner. (Stranger) Bruce gets larger and he gets larger than I know his size is, whereas BUNNY didn't. When he stands over in that other corner he looks the same size; I mean, they both look very small. Now he looks larger than he usually does—you know—I mean when we're standing outside here. I notice his feet look larger than they did in the other corner—about size twelves up here and size fours down there. I guess he looks closer, but I can't remember. Yeah he does look closer than he did in the other corner. He got smaller again, but he didn't get smaller until he got practically two steps before the corner. Have him stand in the middle. Now he looks his normal size. Now he looks like BUNNY did when she was way up in that corner.

B. Female, married 11 months: DAVID? DAVID looks small too. He doesn't look—uh—really as small as Bruce did the last time, but he still looks considerably smaller than he is. And Bruce looks the way he did last time. Oh Bruce looks a lot bigger. DAVID'S farther away though. (Change places) Hah. Hah. Hah. DAVID, DAVID—uh—darling you look pretty much normal except you aren't standing up in the corner as much there so you're closer to the middle of my eyes so you don't look tall. And Bruce looks very small. DAVID looks big.

(Walking from corner to corner—Stranger) Bruce is walking uphill. And now he is very tall, and he was small when he started. Now he's walking downhill, and now he's small again the way he started. (Marital partner) Well DAVID just walked uphill too, but he looks . . . I guess that's all. (He looks what?) Well, he looks more normal to me because I feel I'm a lot closer and it's more in proportion. Oh my, this is horrid. Well now he looks very far away. And he looks smaller, but he doesn't look awfully small. I mean he doesn't look awfully small. I mean he doesn't look—uh—sort of very little. He looks more average size only much farther away.

C. Female, married 4 months: Good night! Well, now the difference is tremendous because Bruce is so small. You see, I can't be objective about it. I know BOB is tall; therefore the difference between his and Bruce's height seems tremendous, more so that the other two. (Does BOB look taller than usual?) Yes, I think. Uh, that's hard to say; he doesn't look taller than usual. Not especially. It's just that Bruce looks so small. He's just minute. That's just about it.

D. Male, married 11 months: Bruce is large. NANCY is small. She is, how you say, Lilliputian. Well as a matter of fact the floor has become, it seems even more slanted. Her feet seem at more of a slant than the floor would indicate. You're (NANCY) just about belt size. Just an armful. You're quite small. When I look at NANCY without looking at Bruce, she looks normal, but when I look at Bruce, NANCY becomes smaller. Nancy is farther away too, and downhill. (Change places) Well, I'll be jiggered. NANCY looks, uh, this way, just looking at NANCY she looks quite normal. Just, I'd say, normal. Bruce looks smaller to NANCY. I think maybe Bruce looks smaller now than he did to Warren, though NANCY is farther away from the ceiling than Warren. Bruce looks very small and downhill too. I guess not as much. The same thing with his feet. His feet seem to be at more of an angle than the floor would indicate. When I look at them both, NANCY looks quite large, though further—no. NANCY looks quite normal actually.

E. Male, married 2 years: SALLY appears about normal to me. Of course Bob looks smaller. But SALLY doesn't look any different. When Bob gets down there and SALLY gets up here it seems to change the appearance of the room slightly. The back wall there—it seems to make that window next to Bob seem a lot larger to me than the window on the right behind SALLY, a lot bigger than if I were just looking into the room without anyone in it.

F. Male, married 3 months: SHE looks about normal size. They look about equal in height. LOUISE is much—just bigger. HER face is larger, but in height SHE is about equal to Sally. Sally looks about 4/5 her normal size.

G. Male, married 3 months: SHE not only goes from left to right, SHE grows big and SHE also comes a little bit towards me. SHE grows small but it looks so funny to hang on to things. SHE seems to go away from me just a little. Now SHE seems to be down and away from me just a little. Bruce looks the same as last time. When I look between windows he looks like a midget in a small

room. When I look at him he seems like a guy standing in a far corner—not a normal far corner. Jan is very tall. SHE is right up to the ceiling. SHE seems to be more in proportion than Bruce.

EXPERIMENT II

As mentioned previously, this experiment was designed both to confirm the results of the first experiment and to provide a quantitative measure of the phenomenon. A number of assumptions underlie the procedure employed in this experiment. A person walking from the left to the right corner in the large room appears to go from very small, through a point of actual size, to very large. The reverse happens if he goes from the right to the left hand corner. If an *S* is set the task of subjectively determining the point at which the observed individual looks his normal size when walking from either corner, we would expect, on the basis of what was found in the first experiment, that a stranger would have to walk a greater distance from either corner than the marital partner in order to appear normal size. Conversely, since in either corner the marital partner appears normal, or more normal than a stranger, the partner will not have to move as far from that corner as the stranger will in order to appear of normal size. Furthermore, the magnitude of the difference between the distances traversed by the stranger and by the partner can be used as a quantitative measure of the magnitude of the phenomenon. At one extreme would be the case in which both stranger and partner had to walk the same distance; we would assume here that the phenomenon did not occur. The other extreme would be when the partner looked perfectly normal in the corner and did not have to move at all.

Hence the major prediction made in this experiment was that when the *S* is set the above task in the large room the point at which the marital partner appears normal will be significantly closer to the starting corner than will the point at which a stranger appears normal. It was also decided to specify the sex of the stranger, i.e., provide a comparison of the partner with both a stranger of the same sex and a stranger of the opposite sex, in order to see if a sex difference might be observed. This specification as to the sex of the stranger had not been controlled in the first experiment; the strangers used were males.

In this experiment six married couples provided a total of twelve *S*s. All couples but one had been married one year, and that couple had been married one year and three months.

Again, the *S* was first asked for a description of the empty room, observing the room from the specified point. Then an individual (either the partner or stranger) was introduced into the room and stood in one of the two corners. The *S* was asked what this person looked like. If the individual was described as being either too large or too small, *S* then received the following instructions:

When I tell (the individual in the room) to do so, he is going to walk slowly across the back wall. I want you to tell him to stop walking as soon as you feel that he looks his normal size, that is, the same size as if he were standing out here (outside of the room).

After *E* obtained assurance from *S* that he understood the instructions, *S* was asked to make two such judgments, once from the left corner and once from the right, for each of the following:

1. One person in the room only: (*a*) the marital partner, (*b*) stranger of same sex as the subject, and (*c*) stranger of opposite sex to that of *S*.

2. Two people in the room, each one standing in a corner, but only one person moving across the wall at a time: (*a*) the marital partner and stranger of same sex as *S*, (*b*) the marital partner and stranger of opposite sex to that of S, and (*c*) stranger of same sex and stranger of opposite sex.

Each judgment was recorded as the reading of the observed individual's position on a measuring tape which had been placed on the rear wall of the room. The tape read from left to right, and the readings were obtained directly from the tape by noting the point reached by the toe of the advancing foot, regardless of whether the individual had come from the left or right corner.

As soon as the *S* saw his or her marital partner in the room for the first time the experimenter asked for a further description of the room: "Before you go any further, I want you to describe the room to me again in as much detail as you possibly can." This was done to see if the introduction of the marital partner into the room would be accompanied by an increase in the distortion of the room itself as reported by *S*. Both descriptions of the room, before and after the introduction of the marital partner, were recorded on a seven-point distortion scale, which provided one point for each of the following which was described as being distorted in any way whatsoever: the floor, ceiling, walls, windows, doors, furniture, and miscellany (e.g., the lighting fixture).

The readings on the tape measure provided the raw data for the experimental results. The *S* made a total of 18 judgments: 6 each for the marital partner, the stranger of the same sex, and the stranger of the opposite sex. Also available were the before-after distortion scores.

Because of the slant of the back wall it was necessary to convert the measurements in inches on the tape measure into angular displacement. This was done through the following steps:

1. The hypothetical center point on the rear wall was obtained by determining the point on the tape measure which was intersected by the median plane of the observer.

2. All readings were then converted into angular displacement from this center point, all readings to the right of center receiving a plus designation.

3. Reference points of 15° to both the left of the center point and to the right of the center point were selected. All readings coming from the left had a constant of plus 15 added to them; all readings coming from the right were first multiplied by minus 1, and then had a constant of plus 15 added to them. Consequently all minus numbers

were cancelled out and comparable distance scores were obtained for all *Ss* coming from either corner. It should be pointed out here that, in terms of the original hypothesis, the actual distance traversed by each individual was not of primary interest. It was the difference between the points reached by different individuals that was of interest, and this difference remains unchanged no matter where the reference point is selected. The reference point is used for computational convenience.

4. For each individual observed by an *S* a mean position score was obtained by averaging the judgments (transformed as described above).

RESULTS

The results for the experiment are presented in Table 2. An examination of this table indicates that the basic hypothesis is borne out: the marital partner was required to walk a shorter distance to be judged his or her normal size than either of the two strangers. The difference of 1.45° between the marital partner and the stranger of the opposite sex is significant at the .001 level of significance (t=3.48). The difference of 1.17° between the marital partner and the stranger of the same sex is significant at the .10 level of significance (t=1.95). The difference of .28° between the two groups of strangers is negligible (t=0.83). The differences of 1.45° and 1.17° represent approximately 8 per cent and 7 per cent, respectively, of the observer's visual field within which the differences were observed, and correspond roughly to a six-inch difference on the tape measure.

TABLE 2

DIFFERENCES IN MEAN POSITION SCORES FOR GROUPS OBSERVED MARITAL PARTNERS, STRANGERS OF SEX OPPOSITE TO OBSERVER, AND STRANGERS OF SAME SEX AS OBSERVER

GROUPS COMPARED	RESPECTIVE MEAN POSITION SCORES		*Diff*	*p*
Stranger opposite sex–M.P.	17.09°	15.64°	1.45°	.001
Stranger same sex–M.P.	16.81°	15.64°	1.17°	.10
Stranger opposite sex–stranger same sex	17.09°	16.81°	.28°	.50

An examination of the before and after room distortion scores, obtained as previously described, indicates that the mean room distortion score for the sample of *Ss* increased from 2.5 to 4.0 after the introduction of the marital partner, and this increase is significant at the .001 level of significance. This must be interpreted with care, however, since the nature of the cues offered by the introduction of *anyone* into the room would tend to make the room more distorted, certainly not less distorted. Yet it is interesting to compare the before-after distortion scores for the four *Ss* whose "measured distance" results are most in accord with the basic hypothesis with the four *Ss* whose "measured distance" results are least in accord with the hypothesis. The former group has mean distortion scores of 2.25 and 4.25, a difference of 2.00 in the predicted direction. The latter group has mean scores of 1.75 and 2.50, a difference of only 0.75 in the predicted direction. Naturally such a small sample precludes drawing any definite conclusions from such a comparison, but it is suggestive of the possibility that the room distortion score does have meaning insofar as it aids in selecting those *Ss* whose reported perceptions are most in accord with the basic hypothesis.

DISCUSSION

The phenomenon under study in this paper represents a particular instance of behavior in which an observer reports a lesser degree of distortion of his marital partner than of a stranger, when observed under conditions which normally produce visual distortion, e.g., in the large and the small distorted rooms. This has been termed the "Honi phenomenon." The results from both experiments point to the conclusion that the Honi phenomenon is a reproducible instance of behavior which can be described adequately in both a verbal report and a quantitative measure and which is capable of having the variables associated with it roughly specified. This reported difference in distortion is capable of being quantified to the extent that a measurable difference can be obtained from the points on the back wall at which the marital partner and the stranger are judged to be of normal size when moving

from either corner. In addition to the mere occurrence of the phenomenon, it should be pointed out that within the groups of *S* used marked individual differences were found which apparently were not due to chance factors, but seemed instead to be due to certain variables, as yet unspecified, contained within the situation itself.

The confirmation of the phenomenon plus the likelihood that the observed individual differences are real raise a very fundamental question which this paper cannot and does not attempt to answer. What are the psychological variables which can adequately explain both the phenomenon and the individual variations within the range of behavior provided by the phenomenon? The variable employed in both of the experiments is marriage, or various length of marriage. It may be assumed or inferred that this variable implies certain interpersonal, emotional, and valueful relationships between those married which cannot be assumed or inferred from the relationships of those not married. But the variable is itself a psychologically meaningless one, and no amount of assumption or inference can adequately specify those variables which are psychologically meaningful and which can adequately explain the observed behavior. Such information is obviously not contained in the experiments reported here, and only further experimentation can provide such information.

However, the experiments reported here do provide some hints as to possible hypotheses which can provide a basis for further experimentation. In the first experiment all of the instances of the phenomenon but one were reported by *S*s who had been married less than one year. The differences obtained in the second experiment were from a sample of twelve in which all but two had been married less than one year. This would seem to suggest that the meaningful variables might be found through an examination of the relationships which are typical of the very earliest stages of marital life, although it must be remembered that the phenomenon was first observed in a couple that had been married over 25 years. It would also indicate that any explanation of the results reported here simply in terms of frequency would have to be rejected, since the results indicate a negative correlation between length of marriage and size of the reported or measured difference. In other words, the fact that one person has looked at another person more often does not explain the difference, and the results appear to contradict such an explanation.

Furthermore, it should be pointed out that any explanation of the experimental results must take account of something other, or something more, than just the nature of the retinal pattern and/or the nature of the stimulus configuration. The observer in this situation is faced with a definite conflict; apparently he cannot make both the room and the people in it look normal at the same time. Presumably one of the two has to be seen as distorted. The typical instance is to see the room as normal, and the people as distorted. In the atypical instance, the Honi phenomenon, the reverse appears to hold true: the room is seen as distorted and the marital partner as normal. The fact that both instances can and do occur indicates that in no sense can either instance be considered as "stimulus bound"; the stimulus and its retinal impingement are not, in and of themselves, an adequate explanation of the observed behavior.

It can readily be seen that the work reported here is entirely compatible with a great deal of experimental work which has been done within the framework of the so-called "transactional" viewpoint. As such, this experiment provides another instance in which a truly adequate and satisfactory explanation of the results must come from a consideration of the total "transaction," in which the perceiving mechanism and the stimulus configuration are merely integrally related parts, and in which the assumptions, needs, values, and purposes of the perceiver are equally as important.

SUMMARY

Two experiments were conducted in an attempt to see if an observable and measurable difference could be obtained in an observer's description of his marital partner as opposed to his description of a stranger, when both were observed in the large and the small distorted rooms.

The first experiment employed 12 couples, an *N* of 24, and utilized a purely verbal method of description. Both the large and small rooms were employed. A number of marked differences were obtained in which the marital partner was reported as being less distorted than the stranger. All but one of these differences were observed with *Ss* who had been married less than one year.

The second experiment attempted to quantify this difference by measuring the difference in the points on the back wall of the large room at which the marital partner and the stranger were judged to be of normal size when walking from either corner. Six couples, an *N* of 12, were employed as *Ss*. The obtained difference was significant and showed that the marital partner moved a shorter ditsance from either corner than the stranger to be judged normal.

It was pointed out that the variable employed in this experiment, marriage or length of marriage, was psychologically meaningless and could only suggest hypotheses for further experimentation which might provide a more satisfactory variable or set of variables. It was also noted that any adequate explanation must be provided in terms of something more than just a consideration of the perceiving mechanism and the stimulus configuration.

Received June 17, 1952. Prior publication.

CASE REPORT

WORKING WITH A PSYCHOPATH

LOUIS C. WEBER
Wichita Falls State Hospital

PSYCHOPATHIC personalities are consistently regarded as a personality type that cannot be effectively rehabilitated; they do not respond to psychotherapy. These immature personalities seem never to adapt themselves to the standards of conduct that guide the normal individual. Theirs seems to be a standard that meets their own personal needs, regardless of the consequences. Not only do they apparently never consider the outcome of their actions, but once committed, they may reenact the same deed knowing full well that such actions will bring punishments previously experienced; but they do not seem to care, or believe that this time it will be different. Some authors attribute this personality type to constitutional factors while others consider it a product of environment. Rodgers (1) hypothesized that psychopathic behavior was an attempt to maintain homeostasis, or a means of controlling anxiety and restoring dynamic equilibrium. Such behavior does seem to be an attempt to gratify certain basic needs.

Few people in the fields of clinical psychology or psychiatry make any determined effort toward restoring the psychopath to society. Such professional persons deem it hopeless; some have remarked that the older the person the more difficult the psychotherapy. But, if a personality diagnosed as psychopathic in early or even late adolescence were given intensive psychotherapy, would it not be possible to develop the superego so necessary to normal personality development? Any adjustment outside an institution would be an asset, even though a few personality defects persisted, provided they were not considered detrimental to the welfare of the community. One theory may postulate that such an undertaking would require the association and identification of the patient with one who is of strong integrity and well-adjusted in his social relationships. Such a "pseudo-father or mother" must possess characteristics that create admiration and affection within the patient. Without these "ties" the patient may well reject the attachment when his needs no longer require the association.

The author, working under the assumption that when a patient diagnosed as "psychopathic personality" seems to possess certain basic personality traits upon which to capitalize, such as the ability to show genuine affection and some identification with his supervisors, believes that this individual may be successfully restored to a normal community life after undergoing psychotherapy and/or counseling of long duration. This report concerns the attempted rehabilitation of such a person, who seemed to be a product of environment rather than a constitutional psychopathic personality, as well as could be determined by the author. A brief outline of the patient's past history will be given along with some results of psychological testing, and then the program of rehabilitation designed to develop a superego and restore him to an active role in society.

The patient, B. T., is a young man, now 21 years of age. He is the youngest son in a family of five children, three of whom are stepsisters by the father. His birth was apparently normal, and he was given much attention and affection by the members of the family. When he was five years of age, he ventured out to the chicken house on the farm where he lived and found his father in a pool of blood surrounded by pigs which were wallowing in the muddy earth. By the father's side lay a shotgun which he had used to commit suicide. The mother refused to accept this verdict and continually referred to it as accidental, indoctrinating B. T. into believing the death was not suicide. Until his commitment to the State Hospital, the patient, when asked about his father, would always reply with, "He died accidentally" or "He was killed."

When he was eight years old, B. T. stole $40 from his grandmother, covered it with paraffin, and buried it in a carefully selected

spot until the missing money was forgotten for about one month. He then removed it from its hiding place and cleaned it, and found it "as good as new." No questions were ever asked concerning the whereabouts of the money, nor was he ever closely questioned regarding the sudden "wealth" he had acquired and spent so freely. There were intermittent periods of petty thievery, hookey-playing, and irresponsible actions throughout his early childhood and adolescent years for which he received little or no punishment.

He started school early, and graduated from high school at 16 years of age, although he was of low average intelligence. He became dissatisfied while attending a junior college and transferred to a senior military college. Here he objected to the rigid discipline of the school and began to fall behind in his studies. By midsemester the patient dropped out of school because of his poor grades. Other than his farmwork as a child, he had been employed on a milk route and at a foundry while he was attending college. He left these jobs because "the work was too hard and the pay too low." Between semesters, when transferring to the senior college, B. T. was hitch-hiking to see his sister and became involved in an automobile-bus accident. He was hospitalized for several days with only minor injuries and later, after his release, he hired a lawyer who helped him obtain a $1500 payment for "injuries" from an insurance company.

After leaving college, B. T. enlisted in the Navy for a one-year period. He felt he could beat the long draft term by this enlistment. During his service, he drank to excess, smoked "reefers," jack-rolled drunks and homosexuals, and would at times engage in promiscuous sexual activities with "pick-ups" in New York City bars. After his one-year service enlistment expired in 1949, he was given an honorable discharge and returned to his home in Texas. However, he was placed in the enlisted reserves for an indefinite period.

When at home, he purchased a car and was employed as an oil field worker. He thought he had contracted syphilis from some girl he had picked up, so he quit his job after one month and went to his grandmother's home where he loafed for two months without bothering to check the supposed disease with a doctor. He then purchased another car, went to a half sister's home and promised to get a job if they would let him live with them. After six weeks more of unemployment, he finally stole his brother-in-law's jacket, went to a large Texas city and obtained employment in an aircraft factory. He rented a room in the home of a divorced woman and became quite friendly with her, although he denies ever having engaged in sexual activities with her. He worked at this job for two or three weeks, on the night shift, and spent his daytime hours visiting around in different towns and with friends, getting little sleep or rest. He quit this job because he felt his employers were about to fire him for sleeping on the job.

Just prior to his commitment to the State Hospital, the patient was working on a plan to bootleg whiskey, and began pawning his property to get spending money. He gave several "hot" checks on himself, talked his half-sister out of $50 even though they had decided to avoid him because of the theft of the jacket, stole a gun from his grandmother and pawned it, then borrowed $300 from her, and purchased various articles "on time." When his mother learned he had taken the gun, and not knowing it had been pawned, she telephoned the police who came to his home and picked him up. While in jail, he said he was cold, so he used the blankets to build a fire in the center of the cell for "warmth." The sheriff became alarmed and contacted B. T.'s mother, and requested she commit him to the State Hospital for observation.

While being transferred to the hospital, the patient was told he would remain for only a few hours and then be released. After his commitment, he learned he would be hospitalized for an indefinite period not to exceed 90 days. With the help of another young patient, he plotted and made good his escape from the hospital the same day he had been committed. He escaped about dusk, parted from his "companion," stole a horse from a nearby farm, and rode the horse to a private flying field a short distance away. He deserted the horse, which he was riding bareback and, convinced that he could fly the plane (although he had never flown before),

he started the motor and began to take off. Halfway down the runway the plane nosed over. The patient escaped with minor injuries and hailed a ride with a motorist on a nearby highway, with the airport authorities in close pursuit.

The motorist who picked him up believed B. T.'s story about having a car wreck and took him into the General Hospital. B. T. stayed there overnight without being identified and, becoming dissatisfied with the food, reported that he escaped from the State Hospital so that he would be returned where "the food is a hell of a lot better."

During his first few days on the ward, the patient was seen several times by the author who observed the patient boasting of his deeds and acting with a superior manner and attitude. When given a newspaper clipping reporting his theft of the plane, B. T. read it and reread it several times, and then boastfully displayed it to the other patients on the ward saying, "That's the plane I stole, fellas; it takes guts to do a thing like that."

For the next few weeks, during routine laboratory and physical examinations, B. T. continued to capitalize on his achievements, boasting of the "big" things he did on the outside and how he had set fire to the jail causing a great commotion. He was transferred to another ward a short time later, and while being visited by his family he boasted of how easy it was for him to break out. The family became alarmed and reported his threats with a request that he be placed under more careful surveillance. B. T. was transferred back to his former ward and placed in a locked room for the remainder of his stay in the hospital.

At first he would talk his way out of working on the ward, or encourage the attendants to give him "rest" periods out of his locked room. The author and hospital social worker visited B. T. frequently, attempting to win his confidence; each time he encouraged his "visitors" to bring him books, cigarettes, model airplane kit, or some other treat. He was evasive and continually on his guard, but never lied to the social worker or author, although he did to others who talked with him. His letters were full of distrust and contempt for his relatives because of his commitment, cursing and blaming everyone he wrote.

On one or two occasions, claiming he was cold in his hospital room, B. T. started a small fire in a waste can. He became involved with another psychopath in the hospital and was able to persuade him to slip him cigarettes, matches, candy, and other things he wanted but had been denied. On one occasion, he had convinced the writer that he wanted to make model airplanes to help pass the time away. A kit was purchased and given to him in hopes that it would help to win him over, and keep him occupied so that he would not be thinking of some sort of trouble to start. The next day the author went down to the ward to see how the project was progressing, only to find that there were two others helping him with the model making. One was his psychopath friend, and the other a homosexual who had been induced to purchase cokes for the other two as a bribe to allow him to work on the plane. B. T. told the author that more glue was needed to complete the plane, although the glue supplied them was sufficient to complete the plane according to the manufacturer's specifications. They had only one wing glued, yet all the glue was gone. After questioning, the psychopath friend admitted that they had put the glue in their cokes because they got a "buzz" out of it. No further glue was supplied, and the plane was never completed.

A series of psychological tests was administered to B. T. during his stay at the hospital. The Wechsler-Bellevue Adult Intelligence Scale, Form I, indicated he was of dull normal intelligence; the Bender Gestalt Test showed no impairment, but did indicate artistic ability and good planning abilities. The Wechsler Memory Scale was in keeping with his intelligence level, and the Minnesota Multiphasic Personality Inventory was invalidated by the high Lie score obtained.

The Thematic Apperception Test was also administered with the following interpretation summary of the results:

> In this test the patient reveals a somewhat unresolved oedipus complex with intermittent "spurts" of ambivalence. He expresses a desire to unbind himself from social convention and submit to his emotional whims and fancy. There is an indication of a fixation at a childhood level, and he feels an

injustice was done him during this period. The desire for "easy money" and immediate gratification of his needs and wishes reflects an attempt to make up for his "lackings" in his formative years. There is a sense of guilt that has not been entirely obliterated from his thinking.

The Rorschach test shows some stereotypy and a failing to identify with the group. He reveals a rather strong ego, but lacks a creative pattern of his own, capitalizing upon those already established. He follows his emotional dictates, being somewhat suggestible yet easily irritated. He appears to be more ego-centered than allo-centered, and he has not made a mature adjustment to people in his environment. There is a noted absence of human content, and a maladjustment to the human element in life.

At a later date, Form II of the Wechsler-Bellevue Scale was administered, in addition to a retest on the Rorschach test. This was after the patient had been placed in semi-isolation (locked in a private room with only an occasional visitor), and a close, friendly association had been established between B. T. and the writer. Here a considerable increase in his intelligence quotient was noted, with a high average level indicated. The Rorschach was essentially the same as on the previous testing with the exception that there was a notable increase in depression, and the development of a stubbornness and uncooperativeness. This testing gave some evidence of a schizoid personality that was not evidenced in the first testing.[1]

It became regular routine for the author to visit B. T. two or three times weekly, when off duty, to talk with him, listen to his stories, and help point out the invalid concepts he had. When the author became too involved in other matters to visit with the patient, the social worker assigned to the case was requested to substitute. She would bring literature for him to read, and would often consult with him in a motherly fashion. It would seem that the social worker was able to reach the patient more effectively at these times, as such statements as "Don't pay any attention to what that guy Weber tells you when you come to visit me. He thinks he knows it all, but someday I'm going to beat him at his own game" were often found in B. T.'s outgoing mail.

Much of B. T.'s mail could not be sent out since he used strong profanity when discussing the relatives who were responsible for his hospitalization, and he often made statements that the hospital staff felt he would regret upon his release. When he wrote to old girl friends, he would propose to each one and give each to understand that she was the only one in his life. Ordinarily such letters are discarded, but in this instance the social worker and author saved each letter B. T. wrote that could not be mailed, often taking them to him and discussing the reason why such a letter could not be sent. These letters were then saved until B. T. was released from the hospital.

During one of the discussion periods, the subject of B. T.'s father's death was approached. The author asked B. T. why he always referred to his father as "He was killed," or "He died accidentally." He replied, "What the hell do you want me to say . . . my father committed suicide? That would be fine, wouldn't it? Then everyone would laugh at me and ridicule me." The method of replying to the question "Is your father living?" was discussed, and although he held onto the idea that the death was accidental, future discussions brought out an answer to the question which would not encourage further probing into the death by the general public. Today, when asked "Is your father living?", B. T. will reply, "No, he passed away some years ago," or "No, he's dead." The interview at which B. T. finally accepted the verdict of suicide in his father's death was the only time that the writer had ever seen the patient on the verge of tears. This acceptance followed a sodium amytal interview during which he relived the experience of his father's death. When he recovered from the amytal, he was told what had actually occurred as a child.

After approximately 50 days hospitalization, and with no further treatment than minimal psychotherapy, B. T.'s mother requested his release from the hospital. The ward doctor consulted with the author and, after discussing the matter of continuing the psychotherapy upon his release from the hos-

[1] The author has purposely neglected to elaborate on the psychological tests since the primary purpose of this paper is the attempted rehabilitation through therapy, rather than a personality evaluation through psychometrics.

pital (without fee), it was decided to try him on a furlough; the doctors felt that as soon as the 90-day commitment expired, B. T. would voluntarily discontinue the sessions and return to his former psychopathic behavior. Fees are usually charged patients whether or not they attend the session as an incentive to force their compliance; being on a nonpayment basis, the continuance of psychotherapy by B. T. was deemed improbable. However, the necessary arrangements were made, and B. T. was furloughed on the condition that he meet with the writer at the times designated by him.

It was felt that close contact would necessarily have to be maintained so as to follow B. T.'s thinking and actions without giving him time to revert to his former behavior. Counseling periods were set at three nights a week at the author's home. The patient wanted to take a week of "rest" before seeking employment, and then apply for work and make some effort to repay the numerous debts he had incurred prior to his hospitalization and to pay for the wrecked plane he had used in his escape attempt.

In the company of the author, B. T. went to the airport to contact the plane owner, and arrangements were made to start payments of $25 per month (provided salary permitted) after he had obtained work. B. T. expressed a desire to go to work in the oil fields as money could be earned quickly, and he had no qualifications for anything but labor work.

Regardless of what type employment he undertook, the writer reminded B. T. that time must be allotted for counseling sessions. He was told that these sessions were important to his rehabilitation, and that to miss even one session would permit him to lose ground in his development. While he agreed that he must not miss any, he attempted to get around it by stating that it was more important to earn money to pay off his debts and hot checks, and he thought the counselor somewhat harsh in insisting upon his attendance even if it interfered with his earnings. However, he was finally convinced that from here on his activities were secondary to his counseling sessions.

Two weeks elapsed during which B. T. neither worked nor made an effort to seek work, offering all sorts of excuses to the author. Each excuse was analyzed to show him that he was only procrastinating, and that he must make an effort to find work if he expected to be able to live. His mother had consented to move to the town in which the counselor lived so as to provide a home for her son, but she neglected to state that she was going to make him support her while she stayed at home in leisure, although she was capable of working. B. T. resented this, and offered it as an excuse for him to leave town and seek employment elsewhere. "Besides," he stated, "who will employ me after I've been in the nut house?" The writer asked his cooperation, and stated that he should exhaust the possibilities in his present situation before leaving.

Both the counselor and the patient went to the State Employment Bureau where the case was discussed with one of the vocational placement men who agreed to assist in whatever way possible. The patient was permitted to sit in on the discussion between the writer and the placement advisor so he would not feel that something might be discussed that was harmful to him, or some special plea made in his case. A job as electrician's helper-trainee was secured through the Bureau, for which B. T. was to be paid $40 per week while in training. Unknown to either the counselor or B. T., the work involved driving a delivery truck. The State Laws of Texas prohibit any patient who is committed to a State Hospital the right to drive a vehicle until discharged from the hospital.

While on the job, which he stated he liked quite well, B. T. began boasting of his accomplishments prior to and during his hospitalization, since his employers knew of his commitment. He worked at the job for about 2 weeks, and then the employer telephoned the writer stating he would have to let B. T. go as he needed someone who could drive, and besides, he added, B. T. was disrupting the harmony of his employees with braggadocio. He also reported that articles were beginning to disappear from the shop lately although no evidence pointed toward B. T. (Oddly enough, it was B. T. who found the thief later on while employed in a grocery store. He found the helper who was employed at the shop, stealing from the meat counter.) B. T. was quite upset over the

loss of the job and placed the entire blame upon the counselor; he felt that if the counselor had not informed his employer of his past history he would still have the job. No amount of discussion could convince B. T. that his own bragging actions were, for the most part, responsible for losing the job. However, B. T. resolved that he would find his own employment, and this time he was not going to tell his employer about his hospitalization.

Another two weeks passed and B. T. was still not employed. He was getting very discouraged, although he continued his sessions without fail. He had about decided to leave town when his landlord, from whom he had rented a small three-room apartment for himself and his mother, informed him of an opening for a checker-clerk in a grocery store within walking distance of the house. B. T. applied for and was given the job at the same salary he had been receiving in his first position. He made no mention of having been hospitalized, and he felt that he was now succeeding in fooling people and "pulling the wool over their eyes" regarding his supposed "insanity."

About this time the author noticed B. T. was clinging to him more and more. B. T. had always been told that he could call upon the author at anytime other than his regular sessions if he felt the occasion required it. B. T. began coming to see the writer almost nightly, and even on weekends, on matters of trivial importance. The transference was beginning to reveal itself. The problem now was to transfer it to a more satisfying channel that would develop group affiliations on a more satisfying level than heretofore. However, it was felt that B. T. was not quite ready for group therapy at this time, although the author believes that the rehabilitation will not be accomplished until such therapy is successfully completed. Slavson feels that group therapy is unsuitable for those with a psychopathic personality, particularly where there are others in need of adjustment involved in such group therapies (2). However, this author feels that if such patients are given individual treatment until object relationships and a desire to change are established, then group therapy becomes not only a possibility, but an essential feature to full and complete rehabilitation of the psychopathic personality.

At opportune times (when B. T. complained of his lack of social life) the affiliation with some group such as the YMCA was discussed. The pros and cons of joining the organization were mentioned, with the result that B. T. made a payment on his YMCA membership, unknown to the author. He was not ready for this yet, and probably was pushed into it by the author; B. T. did not attend one meeting, nor did he renew his membership when the next installment came due. The almost continuous association between the patient and counselor was neither discouraged nor encouraged; B. T. was always told that these additional meetings were of his choosing and that he felt a need for the association and discussions, but that he was always welcome. One thing in every discussion provoked B. T.; the author could always point out the illogical points of his thinking and present angles that B. T. had not thought of in connection with his behavior. At one time during the sessions, B. T. made the statement, "Someday I'm going to beat you at your own game, and have the whole thing reasoned out without a flaw." The counselor replied, "When you do that, B. T., there will be no further need for our continuing the counseling sessions. You will have achieved a rehabilitation." The next session was the first session that B. T. missed during the two months he had been following through with his therapy.

This episode seemed to begin the weaning away of B. T. from the counselor. He has not fulfilled that promise he made, but his reasoning and thinking have cleared considerably since the session took place. He began dating one of the female employees at the store, and also "going places" with one of the male employees. He had even remarked, "Gee, I don't even get a kick out of going into a bar for a beer anymore." No mention of his having missed a session was made the first time B. T. reported to his next regular session. It was B. T. who stated that he did not attend because he thought he wasn't getting anywhere with the sessions, and besides, he stated, he had an opportunity to attend church services where there are lots of pretty girls. "You don't want me to give up my

religion because you think I need to see you, do you?" he argued. Again the necessity of strict compliance was discussed, and the excuse analyzed to show its faultiness.

About this time, B. T. came to the counselor gravely concerned over his position at the store. Someone had entered the store who knew he had been hospitalized, and the person was quite friendly with the manager. He was asked what he thought he ought to do about it, and B. T. replied, "I guess the only thing for me to do is quit before I get fired, and leave town." Here, the author felt, was an ideal place for a little role-playing, so B. T. took the role of the store manager while the author was B. T. The very first action the author took was to approach the store manager, in the person of B. T., and the conversation began:

Author: "Mrs. — I would like to have a talk with you in private for a few minutes."

B. T. "All right, but be quick about it as we have lots of customers in the store."

Author: "Have I been a satisfactory worker here?"

B. T. "Yes, you've done everything as we asked you to, and you have not complained about a thing regarding your work."

Author: "Well, since you can find no fault with me, I want to tell you something. I spent two months in the State Hospital, but I am ready to do a decent job of things, and I think there will be no further reason for me to return to the hospital."

This ended the role playing, and B. T. started to leave the home of the counselor, but remarked first, "I get it, she commits herself into saying you are good, and then she can't very well turn around and fire you because you haven't done a good job." He left deep in thought.

The next counseling session revealed a rather jubilant patient. The very first words he uttered were: "I did the same thing that you did the other night, and I'll be damned if she didn't tell me she knew all about it, and that it didn't make a bit of difference as long as I continued to do my job right. Christ, she also knew I came over here to see you . . . you didn't talk with her did you?" B. T. was assured that the counselor had nothing to do with it, and he felt very much relieved and smugly satisfied that he had "accomplished" something.

Meanwhile the 90-day period of hospital commitment had expired and the counselor informed B. T. that this was the crucial period of his rehabilitation; there was no "threat of rehospitalization" over his head to force him to continue the sessions. The decision was now entirely voluntary and would reveal the extent of his sincerity in his rehabilitation. The few sessions following his discharge from the hospital were uneventful, and the author began to sense that B. T. may have been tiring of them and would discontinue them. But, fortunately, he attended one session with the glad news that he was to receive a raise of $5 a week. He felt he earned a raise although he had not expected that amount, and expressed the thought that perhaps these sessions were of some value after all. He also informed the counselor that he was now ready to make a payment on the damaged plane, although for many sessions previous he had expressed a desire to "get out of the deal if there is any way possible." It was at this time that B. T. was advised that two nights a week were all that would be expected of him now.

Several weeks of uneventful sessions passed, and then one night B. T. failed to keep his appointment. During this time there was a gradual diminishing of "extra" sessions, and he seemed to be extending his sources of satisfaction to other individuals; he went fishing with his cousin or neighbor on holidays, and when financially able, he dated in his free time. He paid off all his debts with the exception of the plane, although he was making regular monthly payments of $25 toward it, and even managed to purchase a few shirts which he prized highly. But definite group affiliations still could not be managed. The author discussed the difficulty with the social worker who handled the patient at the hospital, and the missed appointment was discussed. The social worker phoned B. T. and made an appointment to take him to lunch.

B. T. was quite elated that the social worker had not discarded him after his release. He was not only very courteous and gentlemanly with the social worker (something he greatly neglected while at the hospital), but arrived on time for his next appointment with the counselor. He was very apologetic, gave an excuse which he analyzed to show why it was an excuse and not a reason for his missing the session, and

related enthusiastically the events of his luncheon with the social worker. He reported that he had missed his appointment because he had had to take inventory at the store. However, he continued analytically, he really must have not wanted to keep it as he had been told by his employer that he could leave if he had an appointment to keep. If he just told me, he said, that he had to take inventory, he probably could have let it ride as an excuse, but it was not the reason he did not show up. He had dropped by several nights previous to the appointed evening and found me not at home. He was quite disappointed and thought that he wouldn't be there when expected either. This, he concluded, was improper thinking because he had not informed the counselor that he would like to see him that evening, and so should not have felt "let down" when he found the counselor not at home. After two weeks had elapsed, the author informed B. T. that he had progressed sufficiently in his rehabilitation that one session a week would be all that was necessary, but that he was to come at any other time he felt it necessary. Such occasions no longer became necessary.

After six months had elapsed since B. T. was furloughed from the hospital, he had not missed any further appointments and seemed to be making fair progress in his work, although occasionally he expressed some dissatisfaction with the job. It was quickly dispelled in one or two discussion periods, and he was beginning now to look forward to his regular attendance at the Naval Reserve Unit in town, having been informed by Naval authorities that his attendance was mandatory or he would have to return to the service. There was no resentment or bitterness expressed at the "ultimatum" but a thought of making new friends and expanding his social life was expressed. Sessions were gradually diminished until they had reached a minimum of one session every three weeks, but were not limited to a one-hour period. The counselor saw B. T. on occasions other than the appointed times, but these became rather infrequent as compared to his earlier dependence.

Then one day B. T. learned that the counselor was to be married. It was discussed at one of the meetings, with particular reference to his role. He was reassured that there would be no interruption in his appointments, and that he was to continue to feel free to visit when he felt it necessary. During the time the counselor expected to be out of town, no appointments were scheduled, and B. T. was informed that he was to contact the social worker should he feel that he had to discuss some problem. B. T. appeared to accept everything, and was not seen again until the counselor's return to town.

Within a few days, the counselor learned that B. T. had begun to neglect his work during the interim of the last session and the one due that week. He also became somewhat surly, and was finally told that his services were no longer needed at the grocery store. The manager contacted the counselor and stated that she regretted the move and would hire him again if he could assume his former responsibilities as he was well-liked and a hard worker. However, when B. T. came out to the hospital to see the counselor he appeared somewhat depressed and stated that he had another job in mind.

He had again reported without contacting the counselor, and a prior appointment kept the counselor from talking too long with B. T. He did seem to be somewhat more elated when he left, but there was some evidence of concern.

B. T. failed to keep the next appointment. The next day his aunt telephoned the counselor asking if he knew where the patient had gone. It seems that about three days prior to the appointed time, B. T. had taken all his possessions and left town without telling anyone. The aunt stated that he had been talking about California of late, and she decided to write the brother there to see if B. T. had gone to visit him.

Six weeks elapsed before any member of the family made contact with the counselor, and then it was to learn if he had heard from B. T. The patient had been to California, but left within a few days, and had been unheard of since. No further communication was heard on the case until recently, almost a full year since B. T.'s release from the hospital. The hospital social worker met B. T.'s mother at her place of employment and was informed that he had a job and was

doing satisfactorily in another state. The mother did not have many details of his condition or his employment, but knew that he had not been into any trouble and seemed to be keeping this new job without difficulty. She promised to inform the social worker and counselor of B. T.'s address so that some communication could be effected, but she neglected to do so and no further information has been obtained at this writing.

There appears to have been what could have developed into a serious setback in the rehabilitation of this patient; however, the apparent ability to handle his problems on his own in his new location seems to speak for itself as to B. T.'s adjustment. Although the author would hesitate to conclude that there was definite and lasting improvement, it does appear that some improvement took place. In such cases, psychopaths have been known seemingly to adjust over a period of years, only to return to their asocial behavior without apparent cause. That may still be the result in this case, but the groundwork may be sufficient to continue a growth in personality that will permit insight and reasoning to dominate in B. T.'s thinking, and gradually effect a more lasting adjustment.

Psychopaths are not hopeless cases, particularly when youth is in their favor, this author believes. The prognosis may be poor, but painstaking tolerance and counseling (or psychotherapy when applicable) in their behavior difficulties may prove profitable. They must be able to see some immediate goal and be permitted to achieve this goal in the immediate future so that they will feel they are capable of accomplishing things in a more normal way. One success will tend to encourage others. If the psychopathic personality is placed in proper hands while he is still amenable and in his moldable years, we can hope to see a more normal personality as the end product.

REFERENCES

1. RODGERS, T. C. A dynamic study of the so-called psychopathic personality. *J. nerv. ment. Dis.*, 1948, 107, 43–54.
2. SLAVSON, S. R. *Analytic group psychotherapy*. New York: Columbia Univer. Press, 1950.

Received August 1, 1951.

CRITIQUES AND NOTES

ON KEET'S STUDY, "TWO VERBAL TECHNIQUES IN A MINIATURE COUNSELING SITUATION"

REED M. MERRILL
University of Washington

KEET's provocative study (2) stimulated the author to do a similar study. As a preliminary phase of this research, it was decided to replicate the learning experiment in his study. The failure of this replication is reported.

No attempt will be made to review Keet's entire study, but the design of his learning experiment involved the administration of the Jung word-association test. He selected a disturbing word based on the longest reaction time of each subject (*S*), and on an imperfect reproduction of this word upon second administration of the test. This "traumatic" word then became a key word in a verbal serial maze. A second form of the maze was used as the interpolated learning task. Out of 30 *Ss*, 25 forgot the traumatic key word during a subsequent recall period and thus became the *Ss* for the two forms of experimental therapy used in helping them recall the traumatic word.

In our study we duplicated precisely all of Keet's steps, as reported in his monograph, up to the experimental therapy. The 15 *Ss* were college students comparable to Keet's group. The *Ss* were run by two experimenters to control for experimenter bias. There were no significant differences between the results obtained by the two experimenters. We could go no further and compare counseling techniques for three reasons. First, seven of the 15 *Ss* recalled the traumatic word upon return to the first set of key words when, according to Keet, the traumatic word was supposed to be forgotten and the neutral words recalled. Second, there was not the selective forgetting of the traumatic key word. On the return to the first set of key words, the *Ss* forgot the traumatic key word or one of the neutral words in a random fashion. In no case, for the 15 *Ss* in our experiment, was there the perfect recognition and perfect recall of the first two sets of key words to the maze, followed by the forgetting of the traumatic key word and the recall of the neutral words, as occurred in 25 out of the 30 *Ss* in Keet's study. Third, the *Ss* did not manifest the tension which Keet describes as leading to verbalization over the inability to recall and thus providing the motivation for experimental therapy.

Heim (1), in a paper presented to the Western Psychological Association, reported findings similar to these found by the writer and quite different from Keet's.

We see no way to reconcile our findings and those of Keet in the learning situation. Perhaps factors contributed to Keet's results that are not described in his monograph and that make it difficult to duplicate his findings.

REFERENCES

1. HEIM, R. B. An attempt to repeat the Keet counseling-comparison experiment. Paper read at West. Psychol. Ass., San Jose, April, 1951.
2. KEET, C. D. Two verbal techniques in a miniature counseling situation. *Psychol. Monogr.*, 1948, *62*, No. 7 (Whole No. 294).

Received November 12, 1951.

THE EFFECT OF ANONYMITY ON ATTITUDE-QUESTIONNAIRE RESPONSE

PHILIP ASH AND EDWARD ABRAMSON
The Pennsylvania State College

IN a study (1) of the relative effectiveness of several versions of an experimental film in changing attitudes towards Negroes, three attitude scales were employed. The conditions of the research, in which a pretest-posttest design was employed, necessitated name identifications on all the tests. A question was raised, therefore, concerning the extent to which the findings may have been biased in the direction of admitting to more socially approvable attitudes than would have been the case had the questionnaires been completed anonymously. This problem seemed to be important especially in view of protests on the part of a few of the students against signing questionnaires that "might be held against them."

Research on this problem has been limited. In three wartime studies (5, p. 504) of the effect of anonymity in claiming psychosomatic symptoms on the part of pre-induction examinees, small but consistent differences were found between men who remained anonymous and those who signed their

names to psychosomatic complaints questionnaires. The differences, however, were not statistically significant and "were of such a small order as to be of no practical importance." Benson (4) found some evidence that more people give an "undecided" response on candidates for election in an interview situation than by secret ballot. In these studies, the "anonymous" group was a random half of the total sample.

The three tests dealt with in the present paper comprised an ethnocentrism scale, a political-economic conservatism scale, and a scale measuring attitudes towards Negroes. The first two scales were adapted from the ethnocentrism scale and the political-economic conservatism scale developed by Levinson and reported in *The Authoritarian Personality* (2, pp. 142, 158). The third had been developed for the film study (3). In each of the three scales, the items called for a five-choice response (strongly agree, agree, neutral, disagree, strongly disagree) scored 4, 3, 2, 1, 0. The 0 score was assigned to the most "pro" choice in each case.

Students in each of six sociology classes were given the three tests. In three of the classes the tests were given anonymously; in three the students were required to put their names on their answer sheets. There were sixteen students in each of the six groups, a total sample of 96.

The principal findings are reported in Table 1. For none of the scales was the difference in mean score between the anonymous group and the identified group significant. In each case, the variability of the subgroups within the anonymous group and identified group respectively was considerably greater than the anonymous-identified variability. On the Negro prejudice scale the anonymous group had a slightly higher (more prejudiced) mean score than did the identified group. On the other two scales, these relationships were reversed.

Although the sample employed is relatively small, the conclusion seems warranted that the verbally expressed attitudes of college students, as recorded on scales relating to ethnocentrism, political-economic conservatism, and anti-Negro prejudice, are not biased in either a more "pro" or more "anti" direction as a result of the requirement that they sign the scales, thus identifying themselves.

TABLE 1

MEAN SCORES AND ANALYSIS OF VARIANCE FOR COMPARISON OF THE EFFECT OF ANONYMITY WITH THE EFFECT OF NAME IDENTIFICATION ON ATTITUDE SCALE SCORES

		Ethnocentrism Scale		Polit.-Econ. Conservatism Scale		Negro Prejudice Scale	
		Anonymous	Identified	Anonymous	Identified	Anonymous	Identified
Subgroup Means*		37.8	41.9	37.8	39.1	54.1	51.7
		28.9	35.1	38.9	39.1	50.1	49.8
		32.3	28.4	39.1	37.7	49.4	40.3
Group Means		33.0	35.1	38.6	38.7	51.2	47.3
Source of Variation	*df*	Sum of Squares	Mean Square	Sum of Squares	Mean Square	Sum of Squares	Mean Square
Total	95	19851.5	.....	3872.5	....	53914.8	.....
Between Groups	5	2232.4	.....	40.0	....	1787.3	.....
Anon.-Ident.	1	110.4	110.4	.2	.2	375.8	375.8
Groups within Ident.	2	1471.6	735.8†	23.0	11.5	1204.1	602.0
Groups within Anon.	2	650.4	325.2	16.8	8.4	207.4	103.7
Within Groups (Error)	90	17619.1	195.8	3832.5	42.6	52127.5	579.2

* There is no significance to the order in which the subgroup means are listed. The six subgroups of 16 students each were assigned three to each experimental condition.

† *F* ratio significant at the 5 per cent level of confidence. No other *F* ratios approached statistical significance.

REFERENCES

1. ABRAMSON, E. Employment of sound films for restructuring attitudes. Unpublished doctoral dissertation. Univer. of Pennsylvania, 1952.
2. ADORNO, T. W., *et al.* *The authoritarian personality.* New York: Harper, 1950.
3. ASH, P. The development of a scale to measure anti-negro prejudice. *J. soc. Psychol.*, in press.
4. BENSON, L. E. Studies in secret-ballot technique. *Publ. Opin. Quart.*, 1941, **5**, 79–82.
5. STOUFFER, S. A., *et al.* *Measurement and prediction.* Princeton, N. J.: Princeton Univer. Press, 1950.

Received November 28, 1951.

EFFECTS OF SEATING ARRANGEMENT ON LEADERLESS GROUP DISCUSSIONS

BERNARD M. BASS AND STANLEY KLUBECK
Louisiana State University

Two suggestions appeared in the literature recently concerning the effects of a spatial factor upon the behavior of individual participation in group discussions. Using a circular seating arrangement, Steinzor (7) found a significant relationship between the seating distance of two discussion participants and the tendency of one to follow the other with a verbal remark. The farther away a given participant was from a speaker the more likely he was to react verbally when the speaker finished.

Since amount of participation and leadership status attained in initially leaderless discussions have been found to be correlated from .65 to .86 (1, 2), from Steinzor's findings one might hypothesize that any participant of a leaderless discussion who had a greater mean seating distance from all the other participants would be likely to attain higher leadership status during an initially leaderless discussion. Thus if, as in various studies reported elsewhere (1, 2, 3, 5 8,), an inverted V-shaped arrangement were used as shown in Figure 1, persons who occupied positions 1 and 7 would have some advantage over persons occupying seat 2, 3, 4, 5 and 6, since the former would have only one participant adjacent to themselves whereas all the rest would have two adjacent. If this were true, some correction or change in seating arrangement would be necessary.

On the other hand, Harris (6) had the impression: "To arrange (the participants) in a semi-circle—the chord of which is occupied by observers —is psychologically bad. The end candidates of the semi-circle are isolated, the central candidate is spot-lighted . . ." (6, p. 75).

From the above, one might infer that the discussion participants who occupied positions 1 and 7 would be at a disadvantage in comparison to those occupying 2, 3, 4, 5 and 6, if their discussion scores depended on observers' ratings.

The purpose of the present study was to analyze the relationships, if any, that existed between the particular seat a participant had during an initially leaderless discussion and the leadership status he attained, as estimated by two observers.

Method

Data on 467 participants of 68 half-hour leaderless group discussions of similar group size were available from a study of 48 mixed college students (4), of 231 advanced ROTC cadets (3), of 48 fraternity pledges (8), and 140 college sorority girls (5). The corrected split-half reliability of rated leadership status attained ranged from .81 to .99. All but one of these studies used an inverted V arrangement with observers sitting at the open end of the V facing the participants. In these studies, seat 1 and seat 6, 7, or 8 were at the ends of the V, depending on whether 6, 7, or 8 subjects were used in a particular discussion, while the other seats were more centrally located and were between two seats. In the Bass and Norton study (4), participants sat on both sides of an oblong table with observers at one end as shown in Figure 2.

Seats 1, 4, 5, and 8 would be likely to attain higher status than seats 2, 3, 6, and 7, assuming that an extrapolation of Steinzor's results was valid. Harris' hypothesis could not be applied in this last cited study.

Results

Table 1 summarizes the relations between seating position and leadership status attained during 68 initially leaderless discussions by a total of 467 participants.

Column 1 indicates the subjects of the particular study from which the data were drawn, column 2 indicates whether an inverted V or a rectangular arrangement was used. The size of a discussion group, the number of groups, and their product—the number of subjects—are displayed in columns 3, 4, 5. The mean leadership status attained by participants of each study in relation to their seating position are shown in columns 6 through 13. Means in bold type are those for end seats; means in regular type face are for those seats more centrally located.

Column 14 indicates the F ratios obtained when the variance of means among seats was compared with the variance of scores for persons occupying the same seat,[1] the error term. For each of the 7 analyses, the F ratios obtained led to the inference that the variations in leadership score from seat to seat were attributable to chance.

A fairer test of the hypotheses suggested by an extrapolation of Steinzor's findings was performed by pooling all the scores of occupants of end seats (seats 1 and 6, 7, or 8 in the inverted V studies and seats 1, 4, 5, and 8 or 1, 3, 4, and 6 in the rectangular arranged discussions) and comparing them with the pooled means based on the scores of occupants of the middle seats. Columns 15 and 16 contrast the mean leadership status of those who occupied an end position with those who occupied a

[1] Another possible source of variance was attributable to variations from discussion to discussion, but since this proved to be extremely small for any particular study, it was pooled with the error variance.

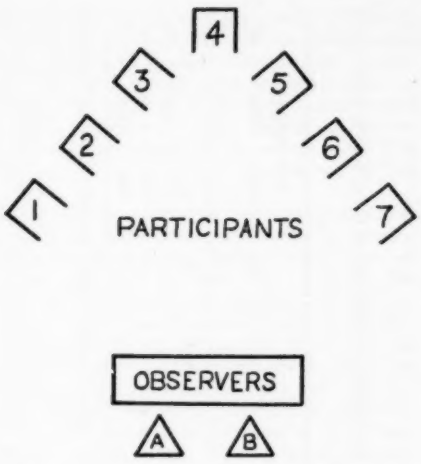

FIG. 1. INVERTED V-SHAPED SEATING ARRANGEMENT

seat between other participants. The last column lists the *F* ratios obtained when the variance of the means between end and middle seats was compared with the variance among persons occupying the same type seat. Among mixed college students meeting 8 per group (Row 1) in a rectangular arrangement, persons occupying end positions attained a mean status more than double that of persons occupying middle seats. Although this difference was significant at the 1 per cent level and although in 5 of the analyses the mean differences were in the direction suggested by the hypothesis originating from Steinzor's study, the variations among individuals occupying a seat of a given type were so great that, despite the fact that the 150 persons occupying end seats on the average attained leadership scores approximately 5 per cent higher than the 317 occupying middle seats, the *F* for the 7 analyses combined was only 1.6—a value to be expected on a chance basis.

Since 467 cases were by no means a small sample, from a descriptive point of view, the investigators regarded the mean difference of 2.2 obtained between persons occupying end seats and those occupying middle seats as a suggestion that a spatial factor was operating in the direction suggested by Steinzor and opposite to that predicted by Harris' hypothesis. However, it was inferred that the presence of many other sources of variance made the rejection of the null hypothesis impossible. For example, there may have been an interaction with group size. Or, the hypothesis might hold for large, rectangular groups but no others. Also it was apparent that for the case of the leaderless group discussions, the factor, if present, was of little practical importance.

One major source of variance which could be eliminated was that attributable to the outside leadership status of the participants. For example, if seat 3 had more than its share of subjects of higher-than-average leadership status outside the discussion situation, because of the correlation ranging from .31 to .47 between discussion and outside status (3, 5, 8), occupants of seat 3 would attain a higher-than-average mean discussion score.

To eliminate this source of variance, analyses of covariance were performed for data based on 275 participants for whom there were available measures of their leadership status or leadership potential outside the discussion situation. These analyses of covariance were performed on the studies whose

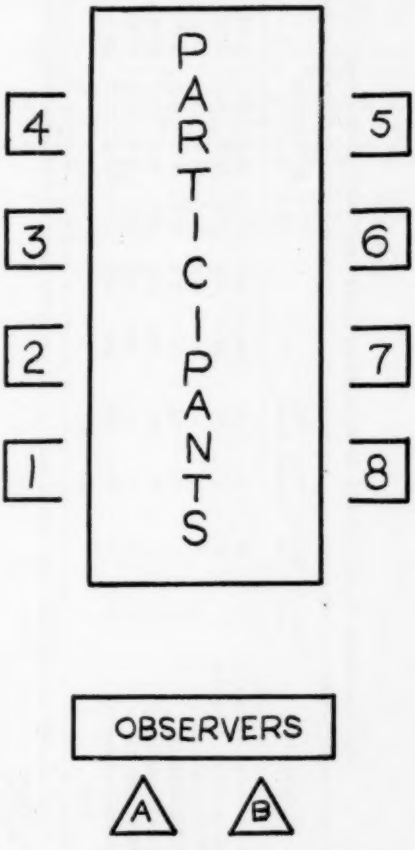

FIG. 2. RECTANGULAR SEATING ARRANGEMENT

TABLE 1

MEAN LEADERSHIP STATUS ATTAINED BY PARTICIPANTS OCCUPYING VARIOUS SEATS DURING INITIALLY LEADERLESS GROUP DISCUSSION

SUBJECTS	ARRANGEMENT	GROUP SIZE	NO. GROUPS	NO. CASES	SEAT NUMBER 1	2	3	4	5	6	7	8	F	SEAT POSITION END	MIDDLE	F
1. Mixed college students (4)	□	8	3	24	40.7†	15.0	30.0	48.7	38.7	23.7	15.0	54.3	1.2	45.6	20.9	8.3*
2. Mixed college students (4)	□	6	4	24	35.0	31.2	50.0	23.8	38.0	38.0			.5	36.7	34.6	.1
3. ROTC cadets (3)	∧	8	12	96	43.1	40.4	39.8	36.0	33.9	42.9	37.6	41.7	.4	42.4	38.4	.8
4. ROTC cadets (3)	∧	7	9	63	25.1	45.4	44.6	36.6	42.2	37.7	47.4		1.8	36.2	41.3	1.0
5. Sorority members (5)	∧	7	20	140	37.3	32.9	38.0	41.5	37.8	39.9	32.8		1.0	35.0	38.1	.1
6. ROTC cadets (3)	∧	6	12	72	48.2	35.7	36.2	44.9	41.4	51.6			1.8	49.9	39.5	6.2**
7. Fraternity pledges (8)	∧	6	8	48	43.4	41.6	37.6	38.4	34.8	43.8			.3	43.6	38.1	1.1
Total			68	467										40.6	38.4	1.6

† Indicates that the 3 individuals who occupied seat 1, an end seat, during one of three discussions attained a mean leadership status of 40.7.
* $p < .01$.
** $p < .05$.

TABLE 2

MEAN LEADERSHIP STATUS ATTAINED ADJUSTED FOR THE EFFECTS OF LEADERSHIP STATUS OUTSIDE THE DISCUSSION OF PARTICIPANTS OCCUPYING VARIOUS SEATS DURING INITIALLY LEADERLESS DISCUSSIONS

SUBJECTS	ORIGINAL GROUP SIZE	CASES WITH OUTSIDE STATUS KNOWN	SEAT NUMBER 1	2	3	4	5	6	7	8	F	SEAT POSITION END	MIDDLE	F
3. ROTC cadets (3)	8	64	−1.4†	.6	4.2	.6	−3.3	.5	1.9	−4.6	.2	−3.0	.7	.57
4. ROTC cadets (3)	7	37	−18.1	1.3	4.6	−.6	10.5	−6.6	1.8		1.5	−8.1	1.6	.22
5. Sorority members (5)	7	140	.8	−1.6	−2.2	2.8	.9	1.8	−2.5		.7	−.9	.9	2.57
6. ROTC cadets (3)	6	34	3.7	−2.0	−7.3	−1.8	−8.7	15.8			1.9	9.8	−5.9	7.17*

† Indicates that the participants of eight-man groups who occupied seat number 1 attained a mean leadership status during the initially leaderless discussions which was 1.4 less than would have been predicted by their average leadership status outside the discussion situation.
* $p < .05$.

data are exhibited in rows 3, 4, 5, and 6 of Table 1. Unfortunately, insufficient data on outside status was available for studies reported in rows 1, 2, and 7. Table 2 shows the mean leadership status attained by occupants of the various seats during initially leaderless discussions adjusted for the discussion status predicted by their outside status. The adjusted means for end seats are displayed in bold type.

Despite the adjustment, the variations among participants could be accounted for by factors other than individual seating position. When the adjusted means for end seats were compared with those of the middle seats, the results were equivocal. Three of the mean differences were in the direction suggested by Harris' suppositions; one was in the direction suggested by Steinzor's results. All of the former could be attributed to chance; the last named was significant at the 5 per cent level of confidence.

It was inferred from these results that, for inverted V arrangements, when the effects of a person's outside status were eliminated, the particular seat a person occupied was of negligible importance in determining a participant's tendencies to attain leadership status during the course of the discussion.

In conclusion, one point must be stressed. The analyses in no way were tests of Steinzor's or Harris' hypotheses. The circular arrangements of Steinzor's and his conclusions were quite different from both the situation and the extrapolation of his results reported in this article. Also, it is quite possible that the opposing factors suggested by both were operating. The isolation of these two opposing effects was not the purpose of these analyses. Rather, it has been shown that if these effects did exist, they tended to cancel each other in initially leaderless group discussions, for it appeared that the particular seat a person occupied had little effect on the final leadership rating he attained during the course of a discussion.

REFERENCES

1. BASS, B. M. An analysis of the leaderless group discussion. *J. appl. Psychol.*, 1949, **33**, 527–533.
2. BASS, B. M. Situational tests: II. Leaderless group discussion variables. *Educ. Psychol. Measmt*, 1951, **11**, 196–207.
3. BASS, B. M., & COATES, C. Forecasting officer potential using the leaderless group discussion. *J. abnorm. soc. Psychol.*, 1952, **47**, 321–325.
4. BASS, B. M., & NORTON, F. T. M. Group size and leaderless group discussions. *J. appl. Psychol.*, 1951, **35**, 397–400.
5. DOLL, P. A. Validity of LGD assessment of unacquainted women. Unpublished master's thesis, Louisiana State Univer., 1952.
6. HARRIS, H. *The group approach to leadership testing*. London: Routledge and Kegan, Paul, 1949.
7. STEINZOR, B. The spatial factor in face to face discussion groups. *J. abnorm. soc. Psychol.*, 1950, **45**, 552–555.
8. WURSTER, C. R., & BASS, B. M. Situational tests: IV. Validity of leaderless group discussions among strangers. *Educ. Psychol. Measmt*, in press.

Received November 1, 1951.

THE ROLE OF ARBITRARINESS IN THE FRUSTRATION-AGGRESSION HYPOTHESIS [1]

NICHOLAS PASTORE

Queens College [2]

SINCE the publication of *Frustration and Aggression* by Dollard, *et al.* (1) there has been a significant increase in the number of experimental studies dealing with reactions to frustrations and an extended application of frustration theory to a variety of problems (2, 3, 4, 5, 6, 7). Of course the book by Dollard and his co-authors and the subsequent work of those who accept the position stated in that book are not alone responsible for the emphasis placed upon the frustration concept, since clinically oriented psychologists have long utilized the concept in their theorizing.

Of the many responses which have been listed as consequences of frustrated acts, aggression is of major importance. This study will be concerned with an analysis of the context of the frustrated act and its relationship to aggression. Although the aggressive response has been singled out for discussion, the opinion of the writer is that the same or a similar analysis applies to the other responses to frustration as well. This point, however, will not be elaborated here.

In a criticism of the frustration-aggression hypothesis the writer pointed out that the reasonableness or unreasonableness of the frustrating agent was significantly related to the evocation of aggressive responses (8). The specific object of this study is to present some evidence for this view.

In an early paper Doob and Sears presented several hypotheses concerning the relationship of substitute and aggressive responses to frustrated acts (2). From their discussion it is evident that they consider aggression to be a primary response to frustration. Whether aggression is expressed behaviorally depends on the relative strength of two main factors: the strength of the goal response and the strength of anticipatory responses for being punished. For instance, for a fixed drive level, the weaker the anticipatory responses for being punished the greater the chance that aggressive responses will occur. We shall concern ourselves with the aggressive response (whether overt or nonovert). Doob and Sears presented sixteen typical frustrating situations to 185 subjects who were asked to indicate their responses to these situations by checking off one of a set of alternative responses that were given along with each situation. The authors' analysis of the results confirmed their theoretical expectations.

It appeared to the writer that in practically all of the 16 situations used by Doob and Sears the frustrating agent could be described as unreasonable or as arbitrary (8). What would be the result if these situations were modified so that the frustrating agent no longer appeared arbitrary? This question forms the basis of the present study. The expectation advanced here is that such modification in the direction of nonarbitrariness would considerably decrease the number of aggressive responses elicited by the various situations.

METHOD

Two sets of typically frustrating situations, 10 situations in each set, were described verbally.[3] Some items were obtained, with modification, from the Doob and Sears study; others were obtained from the Allport and Allport *A-S Reaction Study;* the remaining few were formulated for this study. One set, referred to as *arbitrary in this discussion,* paralleled the situation in the Doob and Sears study. The other set of 10 items was essentially the same as the *arbitrary* set except that each item was somewhat changed in the direction of nonarbitrariness. It did not seem important to the writer that the Doob and Sears study was not followed in its details since the arbitrary set was intended to reproduce a major aspect of their results.

Although the two sets are labeled as arbitrary and nonarbitrary, they were not so labeled when given to the subjects. Undoubtedly the wording of the items and the situations themselves can be improved upon. The writer expects that essentially the same results would be obtained. The important point concerns the explanation of the differential results. The two sets of situations with the instructions are given in Tables 1 and 2.

Admittedly, a concept such as "arbitrariness" is difficult to define in psychological terms. To the writer the term is related to some dimension of justifiableness. The responses of *Ss* to various situations would then be a function, among other variables, of the justifiableness of the frustrating agent. Operationally the term was defined in accordance with the following procedure. Fifty-two college students

[1] This research paper was read at the December, 1951, meeting of the Institute for Research in Clinical and Child Psychology, New York City.

[2] The writer was at Long Island University when this study was undertaken.

[3] I wish to thank Professor G. Aull of Wagner College for help in phrasing the items and for distributing the questionnaire to her students.

TABLE 1

The "Arbitrary" Set of Frustrating Situations

We would like to get information on your feelings and behavior in certain common situations. There is *no* correct or incorrect answer to any of the items. We are simply interested in knowing how people react. If you were not personally involved in any of the situations, try to imagine what your response would be.

1. You're waiting on the right corner for a bus, and the driver intentionally passes you by.
2. You have heard that an intimate friend spread rumors about you which were unjustified and somewhat uncomplimentary.
3. You left an article of yours in a repair shop. You call for it at the appointed time, but the repair man informs you that he has only just begun to work on it.
4. Your date phones at the last minute and breaks the appointment without an adequate explanation.
5. The clerk in a store where you have been waiting for some time for service purposely disregards you and waits on a customer who has come into the store after you.
6. Your instructor springs an unexpected and difficult examination for which you are poorly prepared.
7. Your neighbor's radio across the courtyard consistently prevents you from falling asleep at night. The neighbor refused to do anything about it.
8. You have been waiting in line for sometime to get into a movie. Someone tries to get ahead of you out of turn.
9. You had to get out of bed after you had gone to sleep in order to answer the telephone. It proved to be a stranger who had gotten the wrong number.
10. You're a private in the army and you apply for a promotion which is denied to you. The promotion is given to a less qualified private who has "pull."

TABLE 2

The "Nonarbitrary" Set of Frustrating Situations

We would like to get information on your feelings and behavior in certain common situations. There is *no* correct or incorrect answer to any of the items. We are simply interested in knowing how people react. If you were not personally involved in any of the situations, try to imagine what your response would be.

1. You're waiting at the right corner for a bus. You notice that it is a special on its way to the garage.
2. You have heard that an intimate friend while drunk spread rumors about you which were unjustified and somewhat uncomplimentary.
3. You left an article of yours in a repair shop. You call for it at the appointed time, but the repair man informs you that it isn't ready because of death in his family.
4. Your date phones at the last minute and breaks the appointment because she (or he) had suddenly become ill.
5. The clerk in a store where you have been waiting for some time for service accidentally overlooks you and waits on a customer who has come into the store after you.
6. Your instructor gives a difficult examination for which you are poorly prepared.
7. Your neighbor's radio across the courtyard consistently prevents you from taking a nap in the late afternoon. The neighbor refused to do anything about it.
8. You have been waiting in line for some time to get into a movie. An old lady tries to get ahead of you out of turn.
9. You had to get out of bed after you had gone to sleep in order to answer the telephone. It proved to be a friend phoning for help in an emergency.
10. You're a private in the army and you apply for a promotion which is denied to you. The promotion is given to a private whose qualifications are superior to yours.

(nonpsychology majors) were asked to evaluate the degree of justifiableness of the situations used in this study. Each *S* evaluated 10 different situations—five arbitrary and five nonarbitrary—which were presented in a randomized order. The instructions were:

Write down within the parentheses the number of the alternative which best describes *your own* evaluation of the actions in each of the ten given situations. Every situation involves some individual whose action frustrates the desire of another. You are to judge the degree to which this action is justifiable.

The alternative judgments were: (*a*) completely justified, (*b*) partly justified, (*c*) partly unjustified, and (*d*) completely unjustified. The meaning of the instructions was made clear with an example. The means of the ratings for each arbitrary situation and its alternative nonarbitrary rewording were computed. Briefly, every arbitrary situation yielded a higher mean rating of unjustifiableness than the corresponding nonarbitrary situation. It may be safely concluded, therefore, that people can discriminate between the two types of situations.

For the next phase of this study it was decided beforehand to use over one hundred *Ss* in order to apply large sample techniques. The two sets of situations were alternately passed out to 131 college students who were selected from three different colleges. They were instructed to write down their responses in booklets provided for this purpose. Generally, the students thought that the situations were quite typical.

RESULTS

The responses to the various situations were independently evaluated by the writer and by a graduate student of psychology. The categories were: aggression (whether overt or nonovert), nonaggression, and nonscorable. The two judges agreed on approximately 80 per cent of the items. Discussion on points of disagreement reduced the percentage of disagreements to 5 per cent. A third judge was consulted on the remaining 5 per cent and his judgments were accepted without discussion. For the 67 *Ss* who were given the arbitrary set, the mean number of aggressive responses was 6.34, with a standard deviation equal to 1.77. For the 64 *Ss* who were given the nonarbitrary set, the mean number of aggressive responses was 3.65, with a standard deviation equal to 1.81. The number of nonscorable items was negligible, forming 5 per cent of the total number of responses and more or less evenly distributed between the two sets of situations. For these data, t was greater than three. Thus the difference between the two sample means is significant at the .01 level and in agreement with the initial expectation.

Repetition of this phase of the study with 47 additional college students yielded completely concordant results. The repetition differed from the first in two ways: (*a*) each *S* evaluated five arbitrary and five nonarbitrary situations in a random order; (*b*) instead of writing down their responses, *Ss* were required to check off one of four alternatives. The four alternatives were: (i) I would feel hostile and would show it in my behavior; (ii) I would feel hostile and would *not* show it in my behavior; (iii) I would not feel any hostility; (iv) I would try to do something about the situation without feeling hostile.

All the items contributed to the differentiation of the two samples in the expected direction. Moreover, it was evident in the written replies of the *Ss* that many of the aggressive responses evoked by the nonarbitrary set were in response to a subjective interpretation of arbitrariness. For example, item 8 (nonarbitrary set) deals with an old lady who breaks into a line out of turn. Many of the *Ss* interpreted this action as an unjustifiable one. Such *Ss* indicated some form of aggression (usually nonovert). Apparently the significance of the differential results would be increased if the *S*'s own evaluation of the situation were obtained first.

DISCUSSION

On the basis of these results the arbitrariness of the frustrating situation forms a significant aspect of the various items. Consideration of the hypotheses set forth by Doob and Sears indicates that it is necessary to include reference to the arbitrary factor. For example, the hypothesis dealing with the direct relationship between evocation of aggressive responses and strength of frustrated drive would hold only if the situation is interpreted by the subjects as arbitrary. The prediction might be made that aggressiveness would bear no relation to drive strength if the subject interprets the situation as a nonarbitrary one. An example from everyday life that seems to exemplify this point concerns the mother who frustrates many of her own desires in caring for her infant without necessarily becoming aggressive. It could be claimed, of course, that aggressiveness is in fact present but is inhibited owing to the character of the situation.[4]

From the point of view of Doob and Sears, it may be argued that we learn to inhibit aggression in situations that are nonarbitrary because society punishes its expression in such situations. On the other hand, we learn to express aggression in arbitrary situations because society does not attach punishment to such expression. Regardless of the final validity of this claim it is apparent that arbitrariness and nonarbitrariness are factors to be considered if we wish to understand and predict the reactions of the individual to specific situations.

It has been suggested that the nonarbitrary situations change the goal responses that are interfered with. For example, the goal responses in situation 1 (arbitrary) might be "wanting to get on a bus; then expecting to get on this bus and not being able to." For situation 1 (nonarbitrary) the goal responses interfered with might be "wanting to get on bus; not expecting to get on this bus because of its sign, which decreases rather than increases the instigation." This type of analysis clearly implies that the changes in the goal responses interfered with depend on the character of the situation. For this study that aspect of the situation that is

[4] On this point, regarding the "damming up" of aggressive impulses and their possible subsequent displacement, there exists no adequate experimental evidence. The writer would reject such a formulation of displaced aggression. However, accepting the framework of current frustration theory for purposes of discussing the obtained result of this study, it can be contended that the displacement of "dammed up" aggression will depend on the arbitrariness of the situation to which aggression appears to have been displaced. Or, in other words, displaced aggression may increase the magnitude of the aggressive response to a given situation but the situation itself must be of a certain character if aggression is to be at all evoked.

critical concerns the degree of justifiableness of the action of the frustrating agent.

Considering the result of this study from a somewhat different point of view, it may be contended that a nonarbitrary situation is not frustrating and therefore does not belong to the same class as the arbitrary situations to which one might expect the evocation of aggressive responses. If frustration is defined in terms of the objective situation, namely, interference with some goal-response (as the term was defined by Doob and Sears), then such an argument cannot be advanced since some goal-response that characterizes the situation is being frustrated, regardless of whether the frustrating agent is arbitrary or nonarbitrary. However, if the definition of frustration in terms of the emotional response evoked in the subject is accepted, then a certain amount of circularity is involved—i.e., a situation is defined as frustrating only when the response of the subject is of a specific kind. In addition, adherence to the latter definition would tend to preclude analysis of the relationship of the character of the interfering agent to the emotional response of the subject.

In conclusion it should be noted that the explanation of the manifold responses to frustrated acts probably will entail analysis of the context of the frustrated act; on the subjective or introspective side, it may entail analysis of the ego. The finding of this study may constitute a step in this direction.

Summary

The investigation reported in this study concerns the partial repetition of a study by Doob and Sears in which the arbitrary character of the situations was unintentionally an unacknowledged factor. Situations (the arbitrary set) were drawn up which paralleled those used by Doob and Sears. Modifications of the arbitrary set in the direction of nonarbitrariness was found to reduce considerably the number of aggressive responses which were elicited. This result was interpreted to imply that the arbitrary character of a frustrating situation should be part of a theory that attempts to explain the relationship between frustration and aggression.

References

1. Dollard, J., Doob, L. W., Sears, R. R., Miller, N. E., & Mowrer, O. H. *Frustration and aggression.* New Haven: Yale Univer. Press, 1939.
2. Doob, L. W., & Sears, R. R. Factors determining substitute behavior and the overt expression of aggression. *J. abnorm. soc. Psychol.*, 1939, **34**, 293–313.
3. Frustration. *Personality*, 1951, **1**, No. 1.
4. Graham, F. K., Charwat, W. A., Honig, A. S., & Weltz, P. C. Aggression as a function of the attack and the attacker. *J. abnorm. soc. Psychol.*, 1951, **46**, 512–520.
5. Himmelweit, H. Frustration and aggression; A review of recent experimental work. In T. H. Pear (Ed.), *Psychological Factors of peace and war.* New York: Philosophical Library, 1950.
6. Maier, N. R. F. *Frustration.* New York: McGraw-Hill, 1949.
7. Miller, N. E., & Bugelski, R. Minor studies of aggression: II. The influence of frustrations imposed by the in-group on attitudes expressed toward out-groups. *J. Psychol.*, 1948, **25**, 437–442.
8. Pastore, N. A neglected factor in the frustration aggression hypothesis: A comment. *J. Psychol.*, 1950, **29**, 271–279.

Received February 18, 1952. Prior publication.

IS THE F SCALE INDIRECT? [1]

THOMAS S. COHN
Santa Barbara College
University of California

Some writers have implied that the more intelligent people can simulate their scores on attitude scales (1, 3). Indirect attitude scales presumably avoid the problem of faking, since the purpose of the attitude measure is concealed from the subjects. An example of this type of measure is the F (predisposition to fascism) scale by Sanford, Adorno, Frenkel-Brunswik and Levinson (1). The F scale has two principal purposes: (*a*) to measure prejudice and (*b*) to measure anti-democratic tendencies at the personality level. Each item was included in the scale on the basis of an hypothesized relationship between prejudice and a hidden meaning embodied in the wording of the item.

Presumably the F scale is indirect; yet, it correlates with intelligence. Levinson (1) reports a correlation of —.48 between the Stanford-Binet and the F scale for a sample of 104 employment service veteran men. The mean Stanford-Binet score for this group was 129.5. Gough (2) mentions a correlation of —.43 between the F scale and the Otis IQ for a sample of 271 high school seniors. The writer found a correlation of —.35 using the Altus Verbal Aptitude Test with 140 college students. In view of the restricted samples used, it is likely that even higher correlations would be expected with more heterogeneous samples.

These data suggest the hypothesis that the more intelligent people are able to penetrate the meaning of the F scale and thereby respond in the "proper" manner. To test this hypothesis, 64 college students were administered the F scale twice. On the first administration they were given the set to respond as would the authors of the F scale, "who are highly educated, intelligent people." For the second administration they were told to respond as unskilled laborers of low intelligence and grammar school education. The mean score for the first administration was 2.39 with an *SD* of 1.48 and the mean score for the second administration was 5.45 with an *SD* of .52. The difference between means was significant at the 1 per cent level, only two individuals failing to respond in the proper direction. The difference in scores between administrations was correlated .44 with scores on the Altus Verbal Aptitude Test. The correlation was significant at the 1 per cent level. It appears that the F scale can be faked and therefore cannot be considered an indirect measure. Furthermore, it would appear that intelligence is related to the ability to "fake good."

[1] This paper was presented in part at Western Psychol. Ass., Fresno, April, 1952.

REFERENCES

1. Adorno, T. W., Frenkel-Brunswik, Else, Levinson, D. J., & Sanford, R. N. *The authoritarian personality.* New York: Harper, 1950.
2. Gough, H. G. Studies of social intolerance: I. Some psychological and sociological correlates of anti-Semitism. *J. soc. Psychol.*, 1951, **33**, 237–246.
3. Kahn, L. A. The organization of attitudes toward the Negro as a function of education. *Psychol. Monogr.*, 1951, **65**, No. 13 (Whole No. 330).

Received June 5, 1952. Prior publication.

PERCEPTUAL DEFENSE OR RESPONSE SUPPRESSION? [1]

EDNA M. WHITTAKER, J. C. GILCHRIST, AND JEAN W. FISCHER
University of Wisconsin

Recently, the adequacy of using word recognition thresholds as the dependent variable in studies of perceptual defense has been questioned (1, 2). In part, the argument advanced against their use runs to the effect that *Ss* may see purportedly to-be-defended-against words as soon as, or sooner than, neutral words but refuse, in certain social situations, to verbalize such recognition until additional confirmation of the percept is forthcoming. Thus, thresholds are raised for tabooed words by a sort of response suppression rather than perceptual defense.

In the course of a recent experiment, the present authors incidentally gathered evidence bearing upon the above issue. Briefly, the experimental setup was as follows. There were three groups of 30 *Ss* each: (*a*) a group of white, highly prejudiced, anti-Negro *Ss*, (*b*) a low prejudice group of white *Ss*, and (*c*) a group of Negro *Ss*. Two female *Es*, one Negro and one white, each ran half the *Ss* in each group on a word recognition series. In this series were four categories of three words each refer-

[1] The experiment reported here is part of a series supported by the Graduate School, University of Wisconsin.

ring to people: positively valued, e.g., "winner"; neutral, e.g., "person"; negative, e.g., "drunkard"; and a category specifically derogatory of Negroes, "nigger," "darky," "coon." Thresholds were determined in illuminance units under a constant time of exposure and all *Ss* were instructed to guess on each trial "even if it's the wildest kind of a guess." Recognition was said to have occurred when two successive correct responses were given and the threshold was taken to be the illuminance level of the first of the two successive correct responses.

The experiment yielded a significant complex interaction, Experimenters × Groups × Value categories, which work now in progress is attempting to unravel.[2] The concern of this note, however, is with the protocols of certain of the Negro *Ss* on the specifically derogatory words. These protocols, taken by the Negro *E*, are not duplicated in the protocols of any other group either with this or the white *E*.[3]

S #NN1. Comments on successive trials with the word "nigger": "That same word isn't coming on each time." "Is that 'nigger' or are those two d's?" "It looks like two g's but I still can't distinguish the first part." "Well, I'm going to do like I did on the rest; I'm going to say 'nigger.'" "'Nigger' . . . but . . . "

S #NN3. The *S* reported voluntarily after reporting the word "nigger" that he saw it "before" he said it but that he wouldn't say it because he "couldn't believe it. Couldn't you tell how weak my answer was? Was that supposed to be a trick or something? You got one of your own people sitting here flashing that."

S #NN8. Comment upon reporting the word "nigger": "Oh! 'Nigger.' I should have gotten that before; I thought that was what it was all the time. I guess I should have gotten that sooner."

S #NN9. Comments on successive flashes of the word "darky": "You have to flash that again; I'm not sure." "'Darky'—yeah—'darky.'"

S #NN11. After reporting the word "nigger," *S* said she saw the word "before" but "didn't want to guess it if it wasn't right." She then expressed the hope that she wasn't "messing up" the data.

After reporting the word "darky," *S* said, "I saw it before I said it; I was just wondering for sure."

S #NN12. Comment after reporting the word "darky": "That's what I am; I can say that since only us folks are in here."

S #NN13. Comment after reporting the word "nigger": "That's what I thought I saw back there. You people are sneaky."

S #NN14. Comments on successive trials with the word "nigger": "Oh! That's a bad word. 'Nigger,' is that what I saw? Look, I didn't see that! Don't put that down or else don't put my name down. Don't connect my name with that in any way; don't even dream it." "I did see it! Hey, woman, what kind of test do you have here? I didn't come down here to be humiliated."

S #NN15. Comments on successive trials with the word "nigger": "Do you know what that looked like?" (Prodded to guess.) "No, I better not say it; maybe I'm wrong. Go ahead." "It looked like 'nigger' but maybe I'm wrong." "M-I-G-G-E-R." "I guess it is 'nigger!'"

It is clear from the protocols that responses were being withheld by some *Ss* until the character of the stimulus was unambiguous. Because this kind of protocol occurred only when Negro *Ss* were interacting with a Negro *E* and then only on the specifically derogatory words, it would also appear that evidence of withholding can be obtained only in unusual circumstances. In light of these facts, it would seem that the use of word recognition thresholds of tabooed words as indexes of hypotheses such as "perceptual defense" is a highly questionable practice. Especially does it seem so when it is pointed out that the word recognition threshold of the Negro *Ss* taken by the Negro *E* on the specifically derogatory category of words did not differ significantly from that of either of the other two groups whose threshold was taken by the Negro *E*.

[2] In general, the interaction seems to indicate an effect we have called "situational set" and to support the Howes and Solomon position.

[3] In some of the protocols a trial or two was given after the threshold had been reached.

REFERENCES

1. HOWES, D. H., & SOLOMON, R. F. A note on McGinnies' "Emotionality and Perceptual Defense." *Psychol. Rev.*, 1950, **57**, 229–234.
2. MCGINNIES, E. Discussion of Howes' and Solomon's note on "Emotionality and Perceptual Defense." *Psychol. Rev.*, 1950, **57**, 235–240.

Received June 16, 1952. Prior publication.

LANGUAGE AND COMMUNICATION. By George A. Miller. New York: McGraw-Hill Book Company, Inc., 1951. Pp. xiv+298. $5.00.

Language has long been an embarrassing topic for psychologists. Everybody agrees about its psychological importance. We repeatedly insist that speech is one of the few distinctively human functions, that language symbolism informs the very texture of thought, and that a very large proportion of significant social interaction takes place through its medium. From outside the psychological camp, the older linguists such as Jespersen and Bloomfield tried to link their discussions to the psychology of their day, and the philosophical analysis of Ogden and Richards and the later semanticists was also psychologically conceived. Yet as psychologists we have scarcely nibbled at the edges of the central linguistic problems. With the significant exception of the study of language development in children, psychological research has not advanced our understanding of language appreciably.

It has not been lack of interest that has discouraged psychological research in this field. To the contrary, witness the vigorous controversy about the role of language in thought, and the amount of effort expended in the genetic study of children's speech (which did not have to await adequate conceptualization of language as a psychological phenomenon). Perhaps it is because we are so immersed in language that it is peculiarly difficult to find a fulcral point from which to come to grips with it. But also, our prevailing ways of formulating our problems may have been deficient in researchability.

Miller's text attempts to provide a researchable approach to language by dealing with language behavior functionally in the context of communication. Here the mathematical theory of information recently developed by communications engineers supplies his underlying orientation. The result is a substantial "recoding" of our information about language and communicative behavior, a recoding that bodes well for further advances in understanding.

The initial chapters concern phonetics and the perception of speech. Here, Miller points out in defense of this arrangement, is where we are on surest ground, and where sound research techniques are most easily demonstrated. The nature of speech sounds in physiological, acoustic, and phonemic terms is clearly presented. Acoustic factors governing the perception of speech are treated in some detail, drawing heavily on studies of the Harvard Psycho-Acoustic Laboratory, of which Miller was until recently a staff member.

Already in these early chapters, the basic notions of information theory are introduced. Working from the mathematical definition of the amount of information conveyed by a message as a function of the number of possible alternatives, Miller shows that a single speech wave can potentially carry vastly more information than is ever actually conveyed in vocal communication. Succeeding chapters bring out vividly how the interdependent probabilities that link the symbols in a sequence of speech drastically reduce the number of alternatives that are actually open. The information conveyed by sequences that follow the morphological and syntactic rules of a given language is a minute fraction of what could be carried by statistically independent symbols. Such *redundancy* affords protection against the errors that would otherwise result from the "noise" factor to be reckoned with in any communication system. Miller's own work with language samples of different known degrees of contextual determination provides an experimental attack on the psychological consequences of redundancy that is able so far to sidestep neatly the problem of meaning.

Miller follows Skinner in interpreting language in the framework of behavioristic association theory: verbal behavior is a form of social behavior in which stimulus-response associations depend upon another organism for their reinforcement. The statistical approach to the study of verbal context, as Miller outlines it, seems highly congruent with an associationist account of speech habits. As Miller points out, however, his is a psychology of the talker rather than of the listener, or, one might add, of the thinker.

Miller then surveys the experimental results on word association, perception, learning, and memory in relation to the influence of verbal habits. Taken in this context, the familiar experimental literature on these topics gains significance from the fresh perspective. Psychologists, it seems, have been studying language behavior without knowing it, just as they have been talking prose. When he comes to problem-solving and thinking, Miller offers "recoding" as a translation for the "restructuring" of the field theorists; up to this point, he finds association theory adequate and more parsimonious. A final chapter brings together some of the very recent social psychological work on communication in small and large groups.

I have tried to indicate above some of the main lines of development in Miller's book; the range of his topics is much wider, including language

development, Flesch readibility counts, content analysis, the influence of mass communications, aphasia, language habits as indicators of personality differences, etc. In fact, one of the serious difficulties of the book is implicit in his statement from the foreword that "The subject of language and communication is at least as broad as the subject of psychology itself" (p. xiii). True enough, the subject is all-pervasive. Yet the attempt to include so much of psychology within less than 300 pages entails a cost in superficiality and in looseness of organization.

The most valuable and stimulating contribution of the book, it seems to me, is the wedding of mathematical information theory and a statisticized association theory to provide a coherent approach to language behavior. The state of the field being what it is, Miller's attempt to achieve a representative coverage of studies pertaining to language and communication for a research-oriented text has somewhat obscured the clarity of his original contribution. But the result has the compensatory advantage for the student of bringing together in new relationship a wealth of relevant research, much of which has hitherto been available only in the technical journals.

The fruits of Miller's proposed wedding of information theory and behaviorism will be interesting to observe. The match is not, it seems to me, an inevitable one. A more cognitive psychology like Bruner's finds information theory equally attractive. An approach to language and communication in such terms might do more justice to the listener, and to the referent of "meaning," without invoking the mentalistic ghosts of which Miller still stands in awe.

This book is not the last word on the psychology of language and communuication. But it is an impressive first word, one that should fill a major gap as a text, and do its part toward stimulating and focusing research in an increasingly active field.

M. BREWSTER SMITH

Vassar College

THE STUDY OF INSTINCT. By N. Tinbergen. Oxford: Oxford University Press, 1951. Pp. xii+228.

American psychology has concentrated more and more during this century on the behavior of mammals, and even among mammals has restricted itself mostly to primates (man, chimpanzee, macaque) and a cheap primate substitute, the rat (which is not studied for its rattiness, so to speak, but as a bearer of anxiety and a maker of cognitive maps). European psychologists in general are even narrower, sticking to human subjects only. Among European biologists however there is a continuing tradition of the study of birds, fish, and invertebrates that would nicely complement psychological studies, if only the two groups could keep in touch. But there has been little such contact. Most psychologists are unfamiliar with the zoological studies of behavior, and it seems equally clear that many of the zoologists are unfamiliar with modern psychology.

The importance of Tinbergen's book is that it presents systematically a large amount of significant work by the European school of biologists, making it accessible to the English-speaking reader. The dominance of learning theory in America, established by Thorndike, Holt, and Watson, makes us persistently forget the constitutional factor in behavior (that is, instinct), and this book may help bring it to the graduate student's attention even if his instructors are by now too set in their ways to give it the attention it deserves.

The book will be, and should be, criticized vigorously. It appears that Tinbergen becomes less competent in his handling of evidence when he gets far from the stickleback and the herring gull, as he does in discussing innate mechanisms in man. In this reviewer's opinion, it is also not possible to divide behavior into two parts, instinctive and learned. There is no behavior independent of heredity, and (I believe) none that is certainly unaffected by learning. Instead of two classes of behavior, what should be kept in mind are two factors determining behavior. Finally, theoretical conceptions that Tinbergen presents have obvious difficulties vis-a-vis current theorizing on this continent. But there are few theoretical systems that do not have obvious defects, and we should at least not make the mistake of assuming that ideas unlike the ones we have chosen to work with are *ipso facto* preposterous.

The style is somewhat academic, but the book is full of fascinating, and solidly authenticated, observations of animal behavior. It should be read by everyone, psychologist or zoologist, who is interested in mechanisms of response.

D. O. HEBB

McGill University

THE BIOLOGY OF HUMAN STARVATION. By Ancel Keys, Josef Brozek, Austin Henschel, Olaf Mickelsen, and Henry Longstreet Taylor, with the assistance of Ernst Simonson, Angie Sturgeon Skinner, and Samuel M. Wells. Minneapolis: University of Minnesota Press, 1950. Vol. I, pp. xxii+761, Vol. II, pp. viii+767 to 1385. $24.00.

This published report of critical historical reviews and of original many-sided research on the relation of food deprivation to human behavior and performance is one of the major scientific biological contributions that have resulted from World War II.

It represents the labor and creative scientific activity of the staff at the Laboratory of Physiological Hygiene, School of Public Health, University of Minnesota. An indication of the importance of this monograph is the fact that four distinguished specialists outside the Minnesota laboratory have associated their names with the publication by writing forewords for it: Sir Jack Drummond, Director of Research, Boots the Chemists, England; Dr. Russell M. Wilder, Chief of the Department of Medicine, Mayo Clinic; Dr. C. G. King, Scientific Director, The Nutrition Foundation; and Mr. Robert R. Williams, Chairman of the Williams-Waterman Fund.

In the authors' foreword the genesis of this large investigation is explained:

> The need for an inclusive critical treatise on human starvation and undernutrition became abundantly clear to us when, after several years of work on the immediate problems of military subsistence, we began to ponder the large question of the feeding of peoples. Elementary calculations showed that not only were the world's stocks of food dwindling, but also that the war was devouring the resources for food production and distribution. We were astonished to discover the paucity of knowledge on the effects of simple undernutrition, though here surely is the oldest and one of the most persistent of the disabilities which plague mankind.

It is not only true that the world always has its share of hungry people but there is also another aspect, that is, the role played by general or caloric undernutrition in a large portion of all persons who are ill whether during war or peace. Undernutrition whether from primary or secondary causes produces morphological, biochemical, physiological, and psychological changes in human beings such as to make them behave and perform differently from their adequately fed contemporaries. Scientific data delineating the differences between the undernourished and the well-nourished and examining theoretically and experimentally the many facets of this intricate problem have been few and meager.

At the time of World War I the critical need for information concerning human undernutrition was evident and attempts were made to fill in this scientific gap in our knowledge. One of the major studies at that time, 1917, was made at the Carnegie Nutrition Laboratory, Boston, by Dr. F. G. Benedict and associates, of whom the present reviewer had the honor of being one. But this 1917 study was hampered by the exigencies of the war itself. The investigation was too hurried, and adequate baselines before reduction of diet were not attained. The investigators trying to spare the time of subjects, who were college students and anxious to complete their collegiate education before being called into the war, did not make as many kinds of measurements or as many measurements of the functions studied as would have been desirable. The experimental design made no provision for a prolonged after-study of the subjects during rehabilitation. The Carnegie study was in part designed to supply the United States representatives to the Inter-Allied Food Council with useful data, and in part to reveal what might be the result on the civilian population of the United States should we be forced to live at a considerably lower food intake level. These matters, important at the time, were not altogether concordant with a scientifically objective approach to the total problem. And it is regrettable that the report of that investigation was written at a time when pressing needs of the war were so close upon us.

Since 1917 great discoveries and developments have taken place in nutrition, theory and practice, and many scientific measurement techniques are presently available and well-standardized in terms of instruments and methods that had hardly been dreamed of at the time of World War I. Furthermore the whole concept of experimental design, of sampling of biological materials, and of adequate statistical treatment of results has advanced remarkably during the intervening three decades. The staff at the Minnesota Laboratory of Physiological Hygiene, already trained and equipped for research at the frontier of their broad scientific field, were in a uniquely favorable position to undertake basic research on the undernutrition problem in human beings. The investigation was also well-favored in being able to procure a large group of young men who were conscientious objectors to serve as subjects. These men who, for conscience' sake, did not feel at liberty to participate in military warfare, were, however, as consecrated as any soldier could be in their participation in scientific battles for the benefit of human welfare.

Volume I presents the background of the study and covers the morphological, biochemical, and physiological aspects of the investigation. There are chapters on the history of starvation, the sources of evidence and information, the Carnegie Nutrition Laboratory Experiment, and general features of the Minnesota Experiment. The chapters under the division of morphology cover body weight, physical appearance, body fat, morphology of heart, blood vessels and endocrine glands, blood, bone marrow, bones and teeth, skin and hair, and so on. There are nine chapters under the heading of biochemistry that are mainly concerned with aspects of metabolism—protein, nitrogen, vitamin, lipid, mineral, and carbohydrate. Eleven chapters comprise the section on physiology. Functional physiological changes are of course among the first to be discernible by the subject himself and also by his associates. They tend to dominate the problems of the behavior and the survival of the individual. The investigators in this study have paid a great deal of attention to the physiological exploration of starvation. They undertook "observation and

description of behavior and performance, measurement of the separate variables of physiological function, evaluation as to consequences, explanation in terms of immediate mechanism, and, finally, elucidation of the fundamental causes." Food deprivation and shortage has without doubt been a major factor in natural selection and it is to be expected that many of the functional alterations in starvation play useful roles, as adaptations and adjustments, favorable to the survival of the individual. This study has resulted in great scientific advance in the description of major physiological changes associated with semistarvation and subsequent rehabilitation, and in tentative theoretical formulations toward explanation of the biological mechanisms involved. We cannot expect that even the advanced stage of the present science of physiology will yield a final proven explanation of these mechanisms.

The nature of the physiological problems and their peculiar complexity may be illustrated by bradycardia, which appeared in the severely undernourished persons both in the Carnegie Experiment of 1917 and also here in the Minnesota Experiment. The heart rate is abnormally slow in rest and in work, although the relative rise of the rate in work may be normal. There is a decrease in cycle-variability, both absolute and relative, and extra systoles are very rare in the starving men. This may mean a decreased excitability in the ventricals and that the reduced work of the heart operates to protect diminished cardiac reserve. The decrease in metabolism of the body here seems to be matched by a corresponding reduction in circulatory demand and thus a slower heart rate. But this does not explain the mechanism of producing the slower heart rate. In these experiments the minimal heart rate appeared fairly early in the starvation process when the body weight had been reduced by 15 per cent; thereafter the heart rate gradually increases. This may mean that the heart cannot compensate for the physiological situation by slowing its rate after it itself undergoes more profound degeneration. In realimentation with abundant food the rise in heart rate may reach a level that surpasses that of the prestarvation period. How to evaluate these heart-rate changes in relation to starvation and recovery is thus evidently a complex question.

Most chapters throughout both volumes begin with an historical critical review of pertinent data from the scientific and quasi-scientific literature extending from the remote past to the current period. A great deal has been written about humanity's struggle with famines and food hunger. The bibliography covers 87 pages, amounting approximately to 2300 items. It is indicative of the breadth and significance of this report that in a prefatory note for the section on references the authors are able to say: "In all but a few rare cases the originals have been consulted." The report therefore stands as the most complete and thorough scientific review of the literature of starvation which has been compiled, and represents the most outstanding and comprehensive laboratory experimentation on starvation in human subjects and rehabilitation from its effects.

Volume II probably has most interest to readers of this JOURNAL, particularly the beginning section entitled "Psychology." This section includes chapters on psychological problems in starvation, behavior and complaints in natural starvation, behavior and complaints in experimental starvation and rehabilitation, intellective functions, personality, psychological case studies, and psychological effects—interpretation and synthesis. The section is prefaced by a quotation from Bertrand Russell, "Those who are hungry have no need of an elaborate philosophy to stimulate or excuse discontent . . . the good is enough to eat and the rest is talk." In the psychological section as in the others, the dominant concern in the Minnesota Experiment was with scientific description and measurement of behavioral events and tendencies rather than with explanation. The scientific temper of this section may be judged by quoting a short paragraph:

> In a few variables, such as performance in intellective tests, there was no important relationship to loss of body weight, within the range of semi-starvation obtained in the Minnesota Experiment. In some characteristics a simple linear relationship was indicated; for example, the changes in the Depression score during the period of nutritional rehabilitation were substantially in direct proportion to the level of caloric supplementation. The mathematical simplicity of such a parallelism should not create an illusion of simple biological relationships. Our understanding of the *intervening* variables, and therefore our capacity to explain the process of recovery in this or any other psychological characteristic, is severely restricted (p. 769).

Psychological and social effects of famine are very pronounced (Chap. 37 and Appendix IV). Those living under such a physiological siege tend to devote the major part of their waking time to seeking food. Many other activities are repressed or eliminated. Sex interest tends to be reduced or to cease, although there is an increase in women offering themselves for prostitution. In both primitive and civilized societies when famine reigns social ties are loosened, the amenities and graces are dispensed with, wives and children are abandoned, homes are deserted. There is a reshuffling in reference to property rights, a great increase in thievery; there is lethargy, weakness, and depression. In the Minnesota Experiment (Chap. 38), the complaints of the subjects were systematically recorded in regularly scheduled interviews conducted by three staff psychologists using standard-

ized forms. "In the self-rating questionnaire each subject rated a series of items such as irritability or moodiness on a 6-point scale extending from 0 (normal or absence) to 5 (extreme)." It was the presence or absence of the complaint that was noted here rather than an attempt at scaling the intensity. The inventory contained 53 items covering aspects of the physiological and psychological condition of the subject. Among the items showing large amounts of deterioration during starvation were "tiredness," "muscle soreness," "irritability," "apathy," "ambition," "self-discipline," and "concentration." (Table 360, p. 822.)

A self-rating system was developed for testing three drives, concern with food, sex, and need for activity, and was used at intervals of four weeks during both starvation and rehabilitation. The plotted curve for food drive shows elevations rising rapidly at first and progressively rising until the end of the starvation period. These elevations correspond to the elevated ratings for appetite and hunger pains. For sex and activity drives the curves show depression in general progressing until the end of the starvation period and then show upward slopes toward the normal. The authors state "The course of the curves conforms closely to the clinical impression; the magnitude of the deviations from the normal appears to have been underestimated by the subjects, especially for the sex drive."

A revelation of the Minnesota Experiment that should be of great importance to clinicians concerns the slow rate of recovery in various physiological and psychological areas after rehabilitation was begun. Data are presented which show that improvements in dizziness and in apathy were most rapid and most complete, while tiredness, decrease of the sex drive, and weakness were slow to improve. The subjects judged themselves to fall short of the prestarvation normal in these respects, even after twelve weeks of the recovery regime. And twelve of the men who rated their degree of rehabilitation at the end of twenty weeks of recovery indicated "a nearly complete return to normality." We cannot generalize these recovery rates and delays; they bear a relationship to the duration and degree of severity of the planned starvation procedures. But they do emphasize the importance of making allowance for the condition of many patients during convalescence. When we sometimes say, "Now you're completely well; forget what has happened and carry on as usual," do we realize our blunder?

Mental preoccupation with food, a characteristic of populations during normal famine, was very prominent in the minds of the men undergoing the Minnesota Experiment. Collecting of recipes and reading of cookbooks were favorite occupations. In bull sessions talk of food tended to dominate the conversation. A few apparently came to resent this "tyranny of food" over their conversation. One expressed disgust over this "animal attitude," and another referred to it as "nutritional masturbation." At the beginning of the experiment the subjects were permitted to chew gum, and chewing two or three sticks at a time—rather quickly replaced by new sticks—became characteristic. One of the subjects is reported to have chewed as many as forty packages per day. According to our Carnegie Experiment, where this habit was also a complication, each stick equalled eight calories. Forty calories to the pack would amount to 1600 calories for 40 packs. This illustrates the stealthy manner in which, for the food-minded, liberties in reference to accessory foods can complicate the experimental situation. After the chewing gum shakedown the Minnesota subjects were restricted to two packages per day.

There are many instances in the report of ways in which the subjects tried to get substitutes for the satisfactions normally derived in eating. Several of the men who had never used coffee or tea took up the use of these, and because some of the men worked up to 15 or more cups a day, a maximum for each subject of nine cups had to be set. Several of the men who had never used tobacco took up smoking. One particularly interesting type of behavior developed, which is referred to as "food souping." At mealtime a man might drink the fluid from his soup and then fill the bowl again with hot water, salt it heavily, and then drink the fluid, and repeat the process two or three times before eating the solid portion of the soup. Large quantities of water were of course consumed between meals.

A variety of changes induced by semistarvation created secondary problems for the experimental subject. His face and body show marks of emaciation, his shoes become too large, his clothes fit poorly, and about the twelfth week of semistarvation edema becomes noticeable in the face, knees, and ankles; the hair falls out in large amounts, and shaving cuts are slower to heal; "physical ability to laugh heartily, sneeze, or blush is reduced or absent during the latter stages of semistarvation." The subjects complained about being cold and having cold hands and feet. In hot summer weather they slept under heavy blankets and wore extra clothing when up and about.

Parts of the report on personality changes might well be made standard supplementary reading for courses in personality. The same may be said for the section entitled "Psychological Case Studies." In fact, the reader must have concluded from this review that the writer regards this work as a veritable mine of information and of great significance to psychiatrists, psychologists, and all who are concerned with psychosomatic medicine. We

may confidently expect that many authors of textbooks during the next ten or twenty years will be clamoring for the privilege of including quotations, data, and figures from this monumental study.

WALTER R. MILES

Yale University

THE LONELY CROWD, A STUDY OF THE CHANGING AMERICAN CHARACTER. By David Riesman in collaboration with Reuel Denney and Nathan Glazer. New Haven: Yale University Press, 1950. Pp. 386. $4.00.

In this stimulating study Riesman uses ideas from many disciplines such as cultural anthropology, population studies, sociology, and psychology to support one brilliant generalization. He focuses the beam of this illuminating insight on a variety of problems to provide a truly scintillating book.

His main thesis is that three periods can be distinguished in the growth of a society, each of which produces a predominant character structure in the members of the society. The first period is one of high growth potential: the birth rate is high, but since the death rate is also high, there is only a slow rise in the census. The second period is one of transitional growth and is said to come about through a decline in the death rate due to such factors as centralization, greater efficiency of the administrative process and growing industrialization. In the third period the rising standard of living leads to a decrease in the birth rate which causes an incipient population decline. The character structure found in the members of a society of high growth potential is termed *tradition-directed*. In the period of transitional growth people have typically *inner-directed* characters, while a society in its third stage develops in its members a character structure which Riesman terms *other-directed*.

The tradition-directed people have little opportunity for individualistic development. Their role in the culture is largely laid out for them, providing emotional security though little freedom. Following Fromm's views Riesman believes that this predetermined role precludes the occurrence of maladjustment, or may even provide some tradition-bound patterns for the occasional deviant. The inner-directed man lives in a period of expansion when tradition can offer less and less direction. Instead, direction is maintained by a built-in gyroscopelike mechanism: the norms and principles are interiorized as the result of deliberate "character training" by rather stern parents. Escape was impossible and the rebellious became innovators or explorers in their hopeless attempts to shake off the ever-present company of the internalized parents. Other-directed man has a radarlike mechanism which constantly keeps him tuned to the peer group on which he is mainly dependent for direction. As proportionally fewer and fewer people work on the land or in the mines and factories, ever greater numbers are engaged in occupations concerned with the service or the manipulation of people. The enforced consideration of others' feelings which is necessary in such a society leaves less and less room for individualistic principles of conduct and produces conformity.

After describing the formation of these three character structures in some detail, Riesman takes up the changes in the roles of various agents of character formation. As parents become less important the teacher takes over part of the task, while peers and the mass media usurp their share. In the following chapters Riesman shows how this theory helps in understanding aspects of such diverse phenomena as progressive education, popularity, fashions, comics, and the mass media. Chapters on the changes in such areas as pleasures, popular culture, success, the meaning of food and sex, and the organization of business conclude the first part of the book.

The second part of the book continues with a discussion of the changes in political styles, political persuasion, images of power, and class structure and compares the contemporary American and Kwakiutl Indian cultures.

So far Riesman has tried to describe without evaluating. It does, however, not take a very sharp ear to detect some overtones of criticism. Riesman apologizes for this in a half-hearted way, and proceeds in the third part of the book to sketch his ideas about possible future developments. These center about one theme. Riesman believes that it is possible and necessary for modern man to develop a degree of autonomy, and sees an opportunity for this in an appropriate use of leisure time. After a critical discussion of mass-production-in-play and of the obstacles to autonomy-in-play he ends with some suggestions: consumership, craftsmanship, and taste exchanging can provide areas of competence for increasing numbers of people who desire some autonomy. In a final chapter "Autonomy and Utopia" Riesman summarizes his hopes for America by insisting that the idea that men are created free and equal does not have to mean that they should "lose their social freedom and individual autonomy in seeking to become like each other."

This third part of the book is somewhat disappointing maybe, but one can hardly blame Riesman for a failure to develop a complete philosophy of tomorrow. More serious seems to me the fact that he avoids a discussion of how this autonomous man would get his direction. Would Riesman favor the growth of an intellectual elite or does he see another solution for the paradox of our democracy? I believe that he has not yet been able to form a

decision, but I enjoyed sharing in his search for one.

The ideas in the book are from time to time supported and illustrated by material from interviews in various communities, but the book is frankly nonexperimental, "verstehend," and exploratory. It provides a wealth of ideas for researchers in many areas. The discussion of autonomy concerns itself with the nature and possibilities of self-actualization, which is the ultimate aim of all psychotherapy. The discussion of the role of comics, mass media, and the teacher should provide valuable clues to anybody working with children, professionally or as a parent. The tongue-in-cheek discussion of modern business in the delightful chapter, "From the Invisible Hand to the Glad Hand" should be of particular interest to social and industrial psychologists.

In the first chapters of the book the theory sets out to embrace Western civilization as a whole, or even all civilizations, but the remainder of the book limits itself to the American scene. It would seem that this is wise since the theory might not hold too well for the European continent beyond certain superficial similarities. Education for instance has not become "progressive" in any European country, even though the ideas originated there. One still finds a widespread respect for authority, particularly in cultural matters, although most countries seem to have passed through their period of transitional growth. It may be they are too interdependent to be regarded as separate societies in the sense of population study. Neither does Riesman take into account the existence of religious groups in which successful control is maintained by mechanisms resembling both tradition and inner-direction.

A final note about Riesman's use of slogans. Riesman has steered clear of flippancy while enhancing the enjoyment of reading by the use of catchy expressions, thus demonstrating that profoundness can be combined with readability. Writers of doctoral theses please note!

STEVEN G. VANDENBERG

University of Michigan

DAS ICH UND DIE REBULATIONEN DES ERLEBNISVORGANGES. By F. S. Rothschild. Basle, Switzerland: S. Karger, Ltd., 1950. Pp. xi+388.

The subject of this book is nothing less than the entirety of psychological functioning in its relationship to the topology and structure of the central nervous system. The foundation is provided by the psychological ideas of Klages and Palagyi to which certain aspects of Freud's metapsychology were joined in an astounding and somewhat forcible synthesis. What the author develops from these premises amounts to a moderate cosmology of its own, distinguished by the breadth of his learning as well as by that of his imagination. The whole vast argument is carried by a Platonic confidence in the power of thought to weld together the distressing fragmentariness of empirical information under the arc of encompassing speculation; and a serenity pervades it for which one would look in vain among the more empirical-minded—scurrying as they are for the facts in the foggy drizzle of uncertainty.

Only an outline of the most elementary propositions of this book is possible here. As the title indicates, its central subject is "the Self and the Regulations of the Process of Experience." Its aim is ". . . to see and understand the connection between mental and somatic events from the point of the mind (*Seele*) . . . and in comparisons between human and animal behavior account for the difference on grounds of man's capacity for Reason." The major aim of the study is, "starting from a phenomenological analysis of psychic events to gain a new understanding of the physiological events related to them, and to comprehend in this way results of Psychology and Physiology through the same hypothesis." The indigenous quality of the human organism consists in an Inner Power of Assertion (*innere Behauptungskraft*) which opposes the simple discharge of vital tensions in order to maintain its psychophysical unity. The drive for assertion works mainly through increase or modification of the static tonus which is thought to be regulated by the cerebellum. Because of the opposition of the original instinctual drives Eros and Thanatos (*Urtriebe*), a polarization takes place in the event of each vital tension (*Erlebnis*) which leads to the organization of a body-related and a space-related pole. Any vital (tension) process has to overcome this polarization in order to run its full course. When, under conditions of intensified assertion, this is not possible, a closure peculiar only to man takes place which leads to an act of consciousness. (One of the difficulties of rendering this schema accurately consists in *Geistigkeit* meaning "rationality" at one time and "consciousness" at another.) Man, in Scheler's words, is the-one-who-can-say-no (*Neinsagenkönner*), the "protestant against all mere reality"—a claim which is certainly substantiated by the course of these speculations.

"The Self (*Ich*) is born at a moment when inner conflict in his vitality bear upon the individual maximally and when the formation of the Self means the salvage of his psycho-physical unity." The emergence of the Self takes place in two steps: in the development of an object opposite the subject, a condition which is regarded as the "Core of the Self," or, as one should perhaps put it, as the center of selfhood. This act leads to a regrouping of the poles of experience (*Seelenpole*) which in turn is the condition of Will and Judgment. Those aspects of the Self concerned with these functions

are defined as *Erscheinung des Ich*, freely translated as Manifest Self. "With the development of the Manifest Self Man first spiritualizes (*vergeistigt*) or rationalizes the reality which he is experiencing in that he (now) grasps entities as objects, phenomena as things, and goals of (his) drives and wishes as purposes." These propositions, or posits, (*Seinssetzungen*) retroactively subject his own nature to the dictates of the principle of Reason which he hed created himself. When the motility of man is not only inhibited but also directed by the conscious posits of the Self, then the Manifest Self is complete. Man has become a person connected with two worlds: the original reality of experiencing vital tensions (*Erlebnis*) and the *Welt des Seins*, meaning the objective world created by consciousness.

While the first part of the book is devoted to the genesis of the Self (an *onthological* genesis rather than an ontho- or phylogenetic one), the second part examines weaknesses and disturbances of the Inner Drive of Assertion among a vast array of phenomena such as play, laughter, the comic, sleep and dream, suggestion and hypnosis, as well as sexuality and some of its disturbances. The third part, in which the author sees his most personal contribution, deals with the nervous and endocrine organization which underlies the psychic events described before in meaningful isomorphism. The fourth and last part is devoted to psychopathology on the general premise that the disturbances of the Core of the Self show themselves in psychosis, those of the Manifest Self in the neuroses.

A critical review of what in essence is a "Metaphysic of the Soul" (Baines) does not seem too fruitful. One may, however, note in passing certain features which seem to be common to many systems of its kind and ideological origin. In German psychology, Scheler became one of its standard bearers, Klages another one. In this pursuit *Anschauung* becomes the method of inquiry. First, a universal schema reveals itself to intuitive awareness that is later filled in with details through a peculiar coupling of deduction and induction. Characteristically, the schema is never derived from specific and systematic observations but from generalized life-wide impression. While many subjects of practical import, such as biology or pathology, are dealt with, the system is essentially concerned with a metamorphosis outside time rather than with anything that happens here and now. Behind the semblance of psychological or physiological processes mythological images become discernible, archetypical models of conflict, of separation and reunion, which seem to trace right back to the peregrinations of a divine substance with which Plotinus and the later Platonists were concerned. Klages' *The Spirit as Adversary of the Soul* is a good example. What should be psychology is rather ontology. As the concern is with an event which is neither historical nor concrete, its subject, "the Soul," is autonomous and affected only by its own internal mitosis. Social learning has no place in this story nor is there much understanding of the relationship of the individual to the group. It is not accidental that *Kultur* means symbolic values, culture a way of living. In spite of a pronounced philosophical attitude which pervades all these speculative anthropologies, from the point of epistemology and a critical theory of science they are quite naive, as if they banked on the good will or on the deference of the reader. In this company beholding is believing. The natural: "Hey, wait a moment, how did you get that?" dies on the tongue when you imagine the author looking at you in dismay, saying "Well if you want to be a spoilsport, go ahead and make your objections." But it cannot be denied, the whole thing is based, for one, on a confusion of various levels or dimensions of reality. The argument oscillates between descriptive data and theoretical constructs as it does between observation and interpretation. Yet while it weaves back and forth between them, the universal schema somehow becomes imbued with a superior kind of reality which make what we can really observe look like pallid epiphenomena. In short, as we are talking about the polarities of the Soul we note one *German* polarity, the abiding influence of the romantic school (Fichte, Schelling, and Hegel), without the all-important correction of the other, the transcendental criticism of Kant.

It will be difficult for most of us to follow the author in his speculations. It would be a great mistake, however, to leave it at that without reminding ourselves that there is still a great deal for us to learn from the approach to which the present author has committed himself. There is a subtlety in these phenomenological analyses from which we could profit. Its authors seem to have a penchant for the complexity of human experience, for symbol, structure, and value which the more evidence-conscious approaches are lacking. They are not so apt to bracket-in whole segments of experience, hoping that they will not be there if we just keep on shutting our eyes to them, as experimental psychology has done. An untrammelled imagination does not only lead to license; there are, amidst all the high-flown speculations, ideas of originality and potential value. A concise summary of this book, if it could be made available in English translation, would still be controversial but it would also be very stimulating.

FREDERICK WYATT

Psychological Clinic,
University of Michigan

GROUPS, LEADERSHIP AND MEN. By Harold Guetzkow. (Ed.) Pittsburgh: Carnegie Press, 1951. Pp. ix+293. $5.00.

Groups, Leadership and Men describes the first five years of cooperation between the Human Relations Branch of the Office of Naval Research and the social scientists through whom it conducts contract research. The papers included are mostly the reports given by the principal investigators at a stock-taking conference called by the Advisory Panel of ONR in 1950. The result is a pace-setting book that provides the reader the opportunity to judge at first-hand the scope, significance, and substantive yield of this federally supported research activity.

Among the most important contributions, it furnishes a variety of insights concerning the conditions under which this government agency stimulates social science research. The sections written by Macmillan and Page as the representatives of the Human Relations and Morale Branch of ONR, by Darley as chairman of its Advisory Panel, and by Guetzkow as the book's editor, all provide information about the philosophy that guides the program. It is clear, for example, that effort is centered on promoting basic rather than applied research. The program is broad, including a wide range of social-psychological research within its scope. Looked at from another angle, the participants in this group are establishing the splendid tradition of developing internal communication through the use of the report conference. Working with the Advisory Panel and with noncontractor specialists from outside ONR, the investigators analyze and evaluate their past work and consider their future course as a group responsibility. Finally by publishing an account like this one, this program is establishing the enviable record of providing the public a full and candid description of its activities. There is no doubt about it, through this activity ONR is contributing in significant measure to the development of important aspects of the social sciences.

By way of introducing the book's content, one can ask a set of questions that also serve to spotlight another characteristic of the Human Relations research program. One can ask, "What selective factors seem to govern approval of the projects on which ONR spends its money? Are certain areas of specialization favored over others? Is one form of 'basic' thinking about a problem preferred to another form of equally 'basic' thinking?" While Darley approaches these questions in the first section by contrasting the proposals approved by the Advisory Panel during five years with those that were rejected, additional answers can be found simply by reading through the rather large number of reports included here. It is quite clear that the papers are not based on an explicit preference for any restricted area of research, nor do they seem to favor a particular version of "basic" thinking. Though specialization may be admired and eclecticism deplored in the case of the individual investigator, one feels it is proper that ONR *not* commit its funds to the support of any one systematic position or to a restricted area during this period of exploration and development. As a consequence of the application of this guiding principle through the period under consideration, the subject-matter presented ranges over the entire domain of social science, and divergent conceptual approaches within the same problem area appear in adjoining sections of the book.

A brief description of the kinds of problems being studied, the institutions where the work is being done, and the investigators responsible for administering the projects will give a picture of the program and the flavor of the book. It should be recognized, however, that in many cases the principal investigator is the administrative head of a team and that much of the creative work is done by specialists whose names are given in the book but are not listed here.

The first division is concerned with groups. Describing work done at the Social Science Research Institute of the University of Illinois, Cattell reports progress made on the problem of establishing dimensions for use in assessing group behavior. The next series of projects are all located at the University of Michigan. Reporting studies conducted at the Research Center for Group Dynamics, Festinger summarizes the several papers that have dealt with the dynamics of social communication in small group settings. French, also representing the Research Center, discusses the problem of identifying and measuring variables that can be used to characterize changes in group productivity. Investigations concentrating attention on the functioning of the decision-making group carried out by the Conference Research Project at Michigan are described by Marquis, Guetzkow, and Heyns. The activities of the Survey Research Center, outlined here by Daniel Katz, are concerned with group problems generally comparable to those described above, though the methods employed by SRC investigators were quite different from them in several cases. More detailed descriptions of several of the Survey Research Center's studies are given by Kahn, Jacobsen, and Nancy Morse, and an overview and a prospectus for this activity of the Center are provided by Angus Campbell. In the final survey in this division, Margaret Mead describes the Columbia University Research in contemporary cultures.

The reports presented in the second section bear more directly on the problem of leadership. Shartle and Stogdill, representing the Personnel Research Board and the Department of Psychology of Ohio State University, outline the ten-year-long program

which has been set up there to develop research methods and to conduct basic investigations in the area of leadership and problems allied to it. Investigations focusing on certain situational and personal factors involved in leadership behavior that were conducted at the University of Rochester are reported by Launor Carter. Completing the presentation of studies concerned with the leadership angle, Fillmore Sanford describes the field studies that relate leadership ideology and preference to follower motivation. The latter were conducted at the Institute for Research in Human Relations in Philadelphia.

A third part offers reports centering more on individuals. This section includes the studies conducted at Swarthmore in which Asch has shown the effects of group pressures on modifying and distorting individual judgments. Explorations into achievement motivation being conducted by the research group at Wesleyan University are reported by McClelland. Investigations relating reasoning and values to language behavior conducted at the University of Maryland are described by Cofer and the studies from the University of Minnesota School of Journalism concerned with the learning of factual information from mass media are outlined by Swanson. Two more papers of a somewhat different nature complete the section. The first describes William Hunt's approach to neuropsychiatric screening and the other details Daniel Harris' analysis of social and psychological factors in rehabilitation of the tuberculous.

In addition to these three divisions Darley contributes a synopsis of the conference. Then Macmillan and Page tackle the problem of discussing the military applications of human relations research in the last section of the book. The appendix supplies a guide for the preparation of ONR research proposals, a glossary of technical terms for the general reader (i.e., variable, correlation coefficient, factor analysis, etc.) and the roster of ONR and the members of the Advisory Panel, past and present. An adequate subject-name index concludes the book.

As might be expected in a program of this breadth, the reports themselves are of uneven quality. While some investigations are models of originality—those described by Festinger, Sanford, and Asch are examples—a few are either based on inferior formulations or else employ rather poor designs. The general level of quality is acceptably high, however, and the trends that can be seen operating as the program develops leave the conviction that its quality will improve steadily in the years to come.

Only one kind of information important for evaluating the whole is not included. It concerns the financial side of the program.

Since this activity is supported from public funds, it would be entirely appropriate for the reader to ask the following kinds of questions: Within what size budget does a program like this one operate? What rules of thumb are applied in weighing the cost of a project against its potential significance when considering whether or not it merits ONR support? How has the availability of funds and the restrictions on the way they can be spent dictated the direction in which the program has developed? What limitations and opportunities should the ONR-sponsored investigator anticipate during the next period? To point this up, take the example of an investigation that must go on for years before it can possibly yield a set of significant answers. Over what time span can ONR funds be committed for the support of such a study? Is sponsorship a short-term proposition, contingent on frequent renewals, or can long-term support—subject to standard precautions—be promised for as long, say, as a decade? It is rather uncommon for research reports to supply a sketch of this general kind, probably because it is recognized that expense per se is a sorry criterion against which to measure research. But the size of a budget and the restrictions on how it can be committed surely do constitute significant reality factors that are heavily weighted in many an administrative decision. It is likely that they are operating here too, although no indication of how they work is given. A picture of this side of such a massive program, with concrete examples of how its course has been governed by the operation of such factors, would help the reader to appreciate some of the advantages and dangers that are intrinsic to sponsored research. It would help professional people to understand the boundaries within which their colleagues who are engaged in directing or conducting government-sponsored research must work. The influences that such factors exert on research activity are undoubtedly great. They should be spelled out in a book that describes an entire program in both retrospect and prospect.

The book will appeal to several audiences. Military people who want readable information about the conduct of social science research will enjoy the straightforward presentation. At a more technical level, students in many different research areas will find a number of helpful general formulations as well as a variety of detailed solutions that may suggest answers for some of their own problems. Finally social scientists and government officials concerned with the way a federal agency goes about sponsoring basic research will find the book most interesting. They will discover what appears to be the kind of happy balance that insures the individual freedom in planning and independence in conducting his research while giving him the opportunity and placing on him the responsibility

for developing and maintaining technical and professional standards.

A few words of appreciation are certainly in order. First the responsible members of the Office of Naval Research should be commended on their foresight in planning and supporting basic research in the behavioral sciences. Second, the social scientists who compose the program should be complimented on the excellent manner in which they are discharging an important public trust. Finally, the editor of the volume should be extended a vote of thanks for keeping the presentation sufficiently simple so that the book will be understandable to a wide audience.

University of Texas ROBERT R. BLAKE

LARGE SCALE RORSCHACH TECHNIQUES. (2nd rev. ed.) By Molly R. Harrower and M. E. Steiner. Springfield, Illinois: Charles C Thomas, 1951. Pp. xx+353. $8.50.

The revised second edition of this book differs from the first in only three major ways. There is a review of new research since the first edition, suggestions for a revision of the Multiple Choice Test as a result of many of the negative findings regarding the early test, and a new section on "card pull" which is a reorganization of previous data broken down by determinants and cards.

In spite of the current general emphasis on cross validation of studies of signs, discriminating items, etc., almost all of the new positive studies reported fail to be concerned with cross validation and results of different studies frequently emphasize different signs. The revision of the Multiple Choice Test may provide a much better instrument, but no convincing evidence of its efficiency has yet been made. Like the original, the present revision provides adequate instructions for the use of the Group Rorschach and Multiple Choice Rorschach and much useful, quantified, normative material.

JULIAN B. ROTTER

Ohio State University

NEW BOOKS RECEIVED

ABRAMSON, H. A. (Ed.) *Problems of consciousness,* Transactions of the 2nd conference, March 1951. New York: Josiah Macy, Jr. Foundation, 1951. Pp. 178. Price $3.25.

BEIER, ERNST GUNTER. *The effect of induced anxiety on flexibility of intellectual functioning.* Psychological Monographs, 1951, Vol. 65, No. 9 (Whole No. 326). Washington, D. C.: American Psychological Association. Pp. 26. Price $1.00.

BOISEN, ANTON T. *The exploration of the inner world.* New York: Harper, 1952. Pp. xiii+322. Price $4.00.

BRODY, EUGENE B., & REDLICH, FREDERICK C. (Eds.) *Psychotherapy with schizophrenics, a symposium.* New York: International Universities Press, 1952. Pp. 246. Price $4.00.

BROWN, HARVEY Q. *The enchanted castle.* Boston: Bruce Humphries, Inc., 1951. Pp. 1038. Price $3.75.

BUHLER, CHARLOTTE, SMITTER, FAITH, RICHARDSON, SYBIL, & BRADSHAW, FRANKLYN. *Childhood problems and the teacher.* New York: Henry Holt, 1952. Pp. xi+372. Price $3.75.

BUHLER, CHARLOTTE, LEFEVER, D. WELTY, KALLSTEDT, FRANCES E., & PEAK, HORACE M. *Development of the basic Rorschach score* (supplementary monograph). Los Angeles: Rorschach Standardization Study (mimeographed), 1952. Pp. iv+71. Price $2.50.

CAMPBELL, CLYDE M. (Ed.) *Practical applications of democratic administration.* New York: Harper, 1952. Pp. x+325. Price $3.00.

CATTELL, RAYMOND B. *Factor analysis.* New York: Harper, 1952. Pp. xiii+462. Price $6.00.

COHEN, FRANK J. *Children in trouble.* New York: W. W. Norton, 1952. Pp. xiv+251. Price $3.50.

CURRAN, CHARLES A. *Counseling in Catholic life and education.* New York: Macmillan, 1952. Pp. xxvi+462. Price $4.50.

FISHER, BURTON R., & WITHEY, STEPHEN B. *Big business as the people see it.* Ann Arbor, Michigan: University of Michigan, Institute for Social Research, 1951. Pp. xii+200.

FRENCH, THOMAS M. *The integration of behavior,* Vol. 1. *Basic postulates.* Chicago: University of Chicago Press, 1952. Pp. xi+272. Price $5.00.

FREUD, SIGMUND. (James Strachey, Trans.) *On dreams.* New York: W. W. Norton, 1952. Pp. 120. Price $2.50.

GLUECK, SHELDON. *Crime and correction: Selected papers.* Cambridge, Mass.: Addison-Wesley Press, 1952. Pp. x+273. Price $3.50.

GOLDSCHMIDT, RICHARD B. *Understanding heredity.* New York: Wiley, 1952. Pp. ix+228. Price $3.75.

GREENACRE, PHYLLIS. *Trauma, growth and personality.* New York: W. W. Norton, 1952. Pp. xii+328. Price $4.50.

HINCKLEY, ROBERT G., & HERMANN, LYDIA. *Group treatment in psychotherapy.* Minneapolis: University of Minnesota Press, 1951. Pp. x+136. Price $3.00.

HORNER, A. MUSGRAVE. *Speech training.* New York: Philosophical Library, 1951. Pp. 176. Price $3.75.

HUMPHREY, GEORGE. *Thinking, an introduction to its experimental psychology.* New York: Wiley, 1951. Pp. xi+331. Price $4.50.

INSTITUT HAVRAIS DE SOCIOLOGIE ECONOMIQUE ET DE PSYCHOLOGIE DES PEUPLES. *Revue de psychologie des peuples.* 7me Annee No. 1, 1er Trimestre 1952.

JEFFRESS, LLOYD A. (Ed.) *Cerebral mechanisms in behavior, the Hixon Symposium.* New York: Wiley, 1951. Pp. xiv+311. Price $6.50.

LEYS, WAYNE A. R. *Ethics for policy decisions.* New York: Prentice-Hall, 1952. Pp. xiii+428. Price $4.75.

LOEWY, HERTA. *The retarded child.* New York: Philosophical Library, 1951. Pp. 160. Price $3.75.

LONDON, LOUIS S. *Dynamic psychiatry, basic principles,* Vol. I. New York: Corinthian Publications, 1952. Pp. vi+98. Price $2.00.

LORAND, SANDOR. (Managing Ed.) *The yearbook of psychoanalysis,* Vol. VII. New York: International Universities Press, 1952. Pp. 271. Price $7.50.

MACDOUGALL, CURTIS D. *Understanding public opinion.* New York: Macmillan, 1952. Pp. xii+698. Price $5.00.

MCKEAN, ELLY. *It's mine!* New York: The Vanguard Press, 1951. Pp. 36. Price $2.00.

MERTON, ROBERT K., GRAY, AILSA P., HOCKEY, BARBARA, & SELVIN, HANAN C. *Reader in bureaucracy.* Glencoe, Illinois: The Free Press, 1952. Pp. 464. Price $5.00.

METTLER, FRED A. (Ed.) *Psychosurgical problems.* New York: Blakiston, 1952. Pp. xii+357. Price $7.00.

MILES, HENRY H. W., COBB, STANLEY, & SHANDS, HARLEY, C. (Eds.) *Case histories in psychosomatic medicine.* New York: W. W. Norton, 1952. Pp. x+306. Price $4.50.

MINER, ROY WALDO. (Ed.) *Transactions of the New York Academy of Sciences.* Baltimore: New York Academy of Sciences, 1951. Pp. 69. Member Distribution.

MUKERJEE, RADHAKAMAL. *The dynamics of morals.* London, England: Macmillan and Co., Ltd., 1951. Pp. xxvii+530. Price 25/—.

NEUROBIOLOGIA. Fundador: Ulysses Pernambucano. 1951, Tomo XIV, Junho, Numero 2, Brazil. Pp. 162.

PALMER, JAMES O. *A dual approach to Rorschach validation: A methodological study.* Psychological Monographs, 1951, 65, No. 8 (Whole No. 325). Washington, D. C.: American Psychological Association. Pp. iii+27. Price $1.00.

PINNER, MAX, & MILLER, BENJAMIN F. (Eds.) *When doctors are patients.* New York: W. W. Norton, 1952. Pp. xix+364. Price $3.95.

POLLAK, OTTO, & COLLABORATORS. *Social science and psychotherapy for children.* New York: Russell Sage Foundation, 1952. Pp. 242. Price $4.00.

POOL, ITHIEL DE SOLA, & OTHERS. *The "Prestige Papers,"* Hoover Institute Studies, January, 1952. Stanford: Stanford University Press, 1952. Pp. vii+146. Price $1.75.

REMMERS, H. H., & BAUERNFEIND, ROBERT H. *SRA Junior Inventory.* Chicago: Science Research Associates, Inc., 1951.

ROMAN, KLARA G. *Handwriting, a key to personality.* New York: Pantheon Books, 1952. Pp. xi+381. Price $6.50.

ROSE, ARNOLD M. *Union solidarity.* Minneapolis: University of Minnesota Press, 1952. Pp. xx+209. Price $3.00.

SALTER, ANDREW. *The case against psychoanalysis.* New York: Henry Holt, 1952. Pp. ix+179. Price $2.50.

SCHEIDLINGER, SAUL. *Psychoanalysis and group behavior.* New York: W. W. Norton, 1952. Pp. xviii+245. Price $3.75.

SCHNEIDERS, ALEXANDER A. *The psychology of adolescence.* Milwaukee: Bruce Publishing Co., 1951. Pp. xii+550. Price $4.00.

SMITH, F. V. *The explanation of human behavior.* New York: Macmillan, 1951. Pp. ix+276. Price $2.75.

SMITH, M. B. *The single woman of today.* New York: Philosophical Library, 1951. Pp. xiv+130. Price $2.75.

STEINFELD, JULIUS I. *Therapeutic studies on psychotics.* Des Plaines, Illinois: Forest Press, 1951. Pp. viii+262.

STONE, CALVIN P. (Ed.) *Annual review of psychology,* Vol. 3, 1952. Stanford: Annual Reviews, Inc., 1952. Pp. ix+462. Price $6.00.

THOMPSON, GEORGE G. *Child psychology.* Boston: Houghton Mifflin, 1952. Pp. xxxiii+667. Price $5.50.

THURSTONE, L. L. *Applications of psychology.* New York: Harper, 1952. Pp. x+209. Price $3.00.

VAUGHAN, WAYLAND F. *Personal and social adjustment.* New York: Odyssey Press, 1952. Pp. xiv+578.

VON STAABS, GERDHILD. *Der Sceno-Test.* Zurich: S. Hirzel Verlag, 1951. Pp. 147.

WINCH, ROBERT F. *The modern family.* New York: Henry Holt, 1952. Pp. xxi+522. Price $3.90.

WOLFF, CHARLOTTE. *The hand in psychological diagnosis.* New York: Philosophical Library, 1952. Pp. xv+218. Price $7.50.

ZBOROWSKI, MARK, & HERZOG, ELIZABETH. *Life is with people (the Jewish Little-Town of Eastern Europe).* New York: International Universities Press, 1952. Pp. 456. Price $5.00.

NEUROLOGY & PSYCHIATRY

Section VIII of Excerpta Medica

The Neurology Section includes abstracts of all articles on neurological surgery in addition to general clinical neuro-physiology.

The section on Psychiatry covers the topics usually contained within the meaning of the word. In addition, the following special branches are dealt with fully:

- ★ Psychoanalysis, individual and analytical psychology
- ★ Sexology, criminology, alcoholism, and drug addiction as related to psychiatry
- ★ Social and industrial psychology and psychiatry, vocational guidance and mental testing
- ★ Psychometrics
- ★ Personality testing and the Rorschach test
- ★ Heredity and statistical studies as they apply to psychiatry
- ★ Mental defect and epilepsy

Annual volume of 1000 pages, including an index classified by author and subject.

Subscription price $22.50 per annual volume

Excerpta Medica Service Corporation
250 Madison Ave.
New York 16, N. Y.
Please enter my subscription to Section VIII, Neurology and Psychiatry, of Excerpta Medica.
I enclose $22.50.

Name .....

Address .....

.....

Prentice-Hall is glad to announce the publication of a significant text in

SOCIAL PSYCHOLOGY

by

SOLOMON E. ASCH

Professor of Psychology, Swarthmore College

This is a provocative and penetrating work by an author who has himself contributed to the field of which he writes. It reexamines some crucial ideas about the nature of man that have guided theory and investigation in an attempt to clarify problems and point the way to advancement. The directness of approach and its suspicion of the "self-evident" recommend it as an introduction for the college student. Because it challenges some fundamental assumptions of contemporary social psychology, this book should also be fruitful reading for mature workers in the field. It will be of particular interest to students for its attempt to extend the concepts of gestalt theory into the region of social facts.

The author has undertaken a close examination of the unique properties of social-psychological phenomena, of the distinctive characteristics of human experience and behavior in society. Therefore his account is neither a collection of undigested data nor an extrapolation of principles of individual psychology to new material. The work also treats of a number of problems which have not usually found a place in discussions of social psychology and it includes substantial reports of the author's hitherto unpublished research.

CONTENTS

1. Introduction: Doctrines of Man
2. Organization in Psychological Processes
3. The Transformation of Man in Society
4. Human Invariance and Change
5. The Problem of Human Interaction
6. Social Facts: The Products of Interaction
7. The Expression of Emotions
8. The Perception of Persons and Groups
9. An Introduction to Group Theory
10. The Ego
11. Social Interest
12. Rules and Values
13. The Fact of Culture and the Problem of Cultural Relativism
14. The Doctrine of Suggestion
15. Structural Factors in the Understanding of Assertions
16. Group Forces in the Modification and Distortion of Judgments
17. The Study of Small Groups
18. Attitudes: Their Sociological Shape
19. Sentiments and Attitudes
20. Some Remarks on Propaganda

Publication date: June 15, 1952

Probable price: $5.50 — 736 pages, 6" x 9"

SEND FOR YOUR COPY TODAY

PRENTICE-HALL, Inc., 70 FIFTH AVE., NEW YORK 11, N. Y.